

# Lecture Notes - Week 1 (Lectures 1-3) <sup>1</sup>

## Statistical Mechanics - Course Summary

1. Thermodynamics (  $\sim 7$  lectures) - macroscopic systems, state variables, laws of thermodynamics, phase equilibria, engines and fridges
2. Equilibrium Classical Statistical Mechanics ( $\sim 7$  lectures) - derivation of thermodynamic quantities, ensembles, entropy of mixing, ideal gas
3. Dynamical Foundations and aspects of non-equilibrium ( $\sim 7$  lectures) - Hamiltonian mechanics, Liouville's equation, Poincaré recurrence, Boltzmann equation and H-theorem
4. Equilibrium Quantum Statistical Mechanics ( $\sim 7$  lectures) - quantum statistics, density matrix, boson and fermion systems, Bose-Einstein condensation
5. Modern Topics ( $\sim 2$  lectures) - simulations, dynamics of fluctuations

Statistical Mechanics is the mathematical theory that underpins our understanding of systems of large ( $\sim 10^{23}$ ) numbers of ‘particles’ (e.g gases, liquids and solids). It turns out that due to the large number of particles, we can make rather a lot of progress in describing the **average** properties of these systems. Such *large* systems are often referred to as **macroscopic**. The ‘particles’ or basic entities which make up these systems are referred to as **microscopic**. Our goal in this course is the study of such large systems which we can approach in 2 ways.

1. We can describe the system by a small number of measurable quantities (**macroscopic** variables, e.g. for a gas pressure,  $p$ , volume,  $V$  and temperature  $T$ ). This approach is known as **thermodynamics**. Thermodynamics makes no assumptions about the microscopic nature of these systems. Rather it is based on general laws which have been obtained empirically (from repeated experiments). Such an approach might seem to lack much explanatory power but this is not the case. These laws provide a coherent logical structure and allow us to classify a wide variety of observations in a way which makes them much easier to understand and even gives us the ability to make predictions. In fact, Albert Einstein said in 1949, “A theory is the more impressive the greater the simplicity of its premises, the more varied the kinds of things that it relates and the more extended the area of its applicability. Therefore classical thermodynamics has made a deep impression on me. It is the only physical theory of universal content which I am convinced, within the areas of the applicability of its basic concepts, will never be overthrown.” Your lecturer concurs.
2. Alternatively we can describe the system **microscopically** with detailed assumptions about its microscopic structure (e.g. the masses, positions and velocities of

---

<sup>1</sup> ©University of Bristol 2012. This material is copyright of the University unless explicitly stated otherwise. It is provided exclusively for educational purposes at the University and is to be downloaded or copied for your private study only.

all the molecules which make up a gas). It should be evident that it is impossible to control experimentally (or keep track of theoretically using mathematics) such details of the system (at least for humans!). For example keeping track of the velocity and positions of the atoms of  $\sim 1$  litre ( $\text{dm}^3$ ) of a pure gas (made up of only 1 type of molecule) at 1 atm ( $10^5\text{Pa}$ ) at very low densities requires that we solve  $\sim 10^{23}$  equations of motion (we would need even more equations at higher densities) which is far, far beyond the capabilities of even the best computers available today. But we are probably not interested in all this redundant information. Rather we are interested in **average** values of functions of the microscopic quantities and we would like to use this information to **predict** the behaviour of macroscopic variables. This is the approach known as **statistical mechanics**. However at first sight this does not seem to simplify the problem or help us much - the most obvious (naive) way to obtain averages is to average over time which *still* requires that we solve all these  $\sim 10^{23}$  equations to know the microscopic quantities at every instant of time. Luckily for us, great progress was made at the end of the 19th century when Josiah Willard Gibbs had the idea that rather than average over time, we should average over independent copies of the system (suitably randomised). Using this insight we **can** obtain expressions for macroscopic quantities like  $p, T$  of a gas (and hence the relationships between them) in what has now come to be known as **equilibrium statistical mechanics**.

More recently with the advent of fast computers, it has become possible to study systems of many particles in the *naive* way discussed earlier by solving directly the equations of motion (i.e. Newton's 2nd law  $\mathbf{f}_i = m\mathbf{a}_i$ ) for each of the  $N$  particles in a system. This is a deterministic approach to studying the problem called **molecular dynamics simulation** and any randomness is due only to the variation in initial conditions. It is time-consuming and computationally expensive and limited by numerical accuracy of the computer. The very largest systems that can currently be studied have  $N \sim 10^6$  particles ( $\ll 10^{23}$  required for macroscopic systems) however it gives some important insights into the problem. Moreover it is the only way to precisely study medium sized systems of say  $\sim 1000$  particles where both thermodynamics and statistical mechanics fail (e.g. the study of the small scale dynamics of small biological molecules).

## Books

1. *Statistical mechanics*, R.K. Pathria, Elsevier 2005, 529 pages.
2. *Introduction to modern statistical mechanics*, D. Chandler, Oxford 1987, 274 pages.
3. *An introduction to chaos in non-equilibrium statistical mechanics*, J.R. Doffman, Cambridge 1999, 287 pages.
4. *Pauli lectures on physics, vol 3 : Thermodynamics and the kinetic theory of gases*, W. Pauli, Dover 2003, 160 pages.
5. *Statistical physics of particles*, M. Kardar, Cambridge 2007, 330 pages
6. *Equilibrium and non-equilibrium statistical thermodynamics*, M. LeBellac, F. Mortesagne and G. Batrouni, Cambridge 2004, 616 pages.

## Some Mathematical Preliminaries

If you are not aware of the following they might be helpful for mastering the material in the course.

### 0.1 Some identities involving partial derivatives

In thermodynamics we often have to consider expressing functions of one particular set of variables in terms of another set of variables meaning one often has to relate partial derivatives with respect to different sets of variables while keeping other sets of variable constant.

Say, we have three variables  $x$ ,  $y$  and  $z$  which are not independent in that  $F(x, y, z) = 0$ . Therefore, we know that:

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \quad \text{and} \\ dy &= \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz. \end{aligned}$$

Substituting the second  $dy$  into the first equation gives:

$$dx = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z dx + \left( \left(\frac{\partial x}{\partial z}\right)_y + \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \right) dz. \quad (0.1)$$

Since  $x$  and  $y$  can be varied independently we have

$$\left(\frac{\partial x}{\partial y}\right)_z = \left[ \left(\frac{\partial y}{\partial x}\right)_z \right]^{-1} \quad \text{and} \quad (0.2)$$

$$-1 = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y. \quad (0.3)$$

Similarly, since  $x$  can be written as a function of  $y$  and  $z$  we have

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz.$$

Dividing by  $dw$  and imposing  $dz = 0$  gives:

$$\left(\frac{\partial x}{\partial w}\right)_z = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial w}\right)_z \quad (0.4)$$

A further useful identity is obtained by letting  $x = x(y, w)$  such that

$$dx = \left(\frac{\partial x}{\partial w}\right)_y dw + \left(\frac{\partial x}{\partial y}\right)_w dy$$

leads to the following by dividing by  $dy$  and imposing a constant  $z$  gives:

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial x}{\partial y}\right)_w + \left(\frac{\partial x}{\partial w}\right)_y \left(\frac{\partial w}{\partial y}\right)_z. \quad (0.5)$$

## 0.2 Some notation : partial derivatives

In these notes a number of different notations for partial derivatives have been used in these notes. As you are no doubt familiar the  $n$ th partial derivative of a function  $f$  w.r.t to a variable  $y$  is given by :

$$\frac{\partial^n f}{\partial y^n} \quad .$$

We will also use the more compact notation

$$\partial_y^n f \quad .$$

## 0.3 Constrained Optimization

**Lagrange multipliers** are often very useful in minimisation (or maximisation) problems where the given system must satisfy certain requirements.

### 0.3.1 Example: Minimum of a function constrained to be on a particular curve on a surface

Imagine that you are asked to determine the minimum of the function  $F(x, y) = x^2 + y^2$  with the condition that the following relationship exists between  $x$  and  $y$ :  $y = 2x + 1$ .

- It is easily solved by expressing  $y$  in terms of  $x$  in  $F(x, y)$ .

$$f(x) \equiv F(x, y(x)) = F(x, 2x + 1) = 5x^2 + 4x + 1 \quad (0.6)$$

$$i.e.: 0 = \left. \frac{df}{dx} \right|_{x_0} = 10x_0 + 4 \quad (0.7)$$

$$i.e.: x_0 = -\frac{2}{5}, y_0 = \frac{1}{5}, F(x_0, y_0) = \frac{1}{5} \quad (0.8)$$

- There is another way to solve this problem. We define a function that depends upon three parameters  $x, y$  and  $\lambda$ :

$$G(x, y, \lambda) \equiv F(x, y) + \lambda(2x + 1 - y). \quad (0.9)$$

A minimum of  $G$  with respect to all three parameters is now determined:

$$0 = \frac{\partial G}{\partial x} = 2x + 2\lambda \quad (0.10)$$

$$0 = \frac{\partial G}{\partial y} = 2y - \lambda \quad (0.11)$$

$$0 = \frac{\partial G}{\partial \lambda} = 2x + 1 - y. \quad (0.12)$$

The solution for this system is then  $y_0 = 1/5, x_0 = -2/5, \lambda_0 = 2/5$  with the same values as before. Note the important property of the definition of  $G$  i.e. that the parameter, the so-called *Lagrange multiplier* appears linearly, in order that the condition for minimization w.r.t.  $\lambda$  then immediately returns the condition  $2x + 1 - y = 0$ . Since  $x$  and  $y$  are not coupled in  $G$  (not the case in  $f(x)$ ) the  $\lambda$ 's second role is to couple the equations for minima in  $x$  and  $y$ .

- But why is it the same solution as in the first case above? We investigate from a third perspective. We vary the function  $F$  as follows:

$$\delta F = \frac{\partial F}{\partial x} \delta x + \frac{\partial F}{\partial y} \delta y. \quad (0.13)$$

But we may only vary it along the prescribed path  $y = 2x + 1$ , and that means that the  $\delta x$  and the  $\delta y$  are coupled. Since we are working with small variations, we immediately know that  $\delta x \propto \delta y$  where we must determine the factor of proportionality. [We ignore terms in *e.g.*  $(\delta x)^2$  or higher order and a relationship of the type  $\delta x = \delta y + c$  would be contradictory to the idea of a continuous path.] Specifically the relationship is:

$$\frac{\partial(y - 2x - 1)}{\partial x} \delta x + \frac{\partial(y - 2x - 1)}{\partial y} \delta y = \delta y - 2\delta x = 0. \quad (0.14)$$

Eqn. (0.14) may now be multiplied with an arbitrary factor  $\lambda$  and added to eqn. (0.13):

$$\delta F = \frac{\partial F}{\partial x} \delta x + \frac{\partial F}{\partial y} \delta y + \lambda(\delta y - 2\delta x) = \left[ \frac{\partial F}{\partial x} - 2\lambda \right] \delta x + \left[ \frac{\partial F}{\partial y} + \lambda \right] \delta y. \quad (0.15)$$

Since the minimisation must take place on the curve  $y = 2x + 1$ , this means that a change in  $x$  must take place at the same time with a change in  $y$  and that

$$\frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{\partial y}{\partial x}. \quad (0.16)$$

But we know that for an extreme value  $\delta F = 0$  and in the handling of eqn. (0.15) the coefficients of  $\delta x$  and  $\delta y$  must both be zero. Hence the second term in (0.15) is

$$\lambda = -\frac{\partial F}{\partial y} \quad (0.17)$$

that yields the coefficients of  $\delta x$  as:

$$\frac{\partial F}{\partial x} - 2\frac{\partial F}{\partial y} = 0 \quad (0.18)$$

and that the correct expression for the extreme value is:

$$\frac{dF}{dx} = \frac{df(x, y(x))}{dx} = \frac{df(x)}{dx} = 0. \quad (0.19)$$

### 0.3.2 Generalization to two and higher dimensions

Consider a differentiable function  $z = z(x, y)$  (we can view it as a surface characterised by hills and valleys). At any point in the area where the function is defined, the gradient is

$$\nabla z = \left( \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right). \quad (0.20)$$

This is a vector in the  $xy$ -surface. Remember that this vector is the direction of the steepest descent of  $z(x, y)$  and that it is perpendicular to the contour line through the point where it is calculated.

Now consider the contour of another function  $g(x, y)$  where  $g(x, y) = 0$ . We now wish to find an extreme value for the function  $F(x, y) = z(x, y)$  only on the curve  $g(x, y) = 0$ . This can only be at a position  $(x_0, y_0)$  where a contour line of  $z(x, y)$  passes through  $(x_0, y_0)$  *tangentially* to the curve  $g(x, y) = 0$ . If these two curves are tangential their gradient vectors must have the following relationship:

$$\nabla z|_{(x_0, y_0)} = \lambda \nabla g|_{(x_0, y_0)}, \quad \lambda \in \mathbb{R}, \lambda \neq 0. \quad (0.21)$$

It follows for a relative extreme value that the following three equations of which the first two derive from the gradient conditions (for  $x$  and  $y$ ) and the third is the prerequisite for the coupling of  $x$  and  $y$ :

$$\frac{\partial F}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \quad (0.22)$$

$$\frac{\partial F}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \quad (0.23)$$

$$g(x, y) = 0 \quad (0.24)$$

This is true, naturally, if the gradient is non-zero.

This argument can be straightforwardly extended to higher dimensions : the extremum of a function  $f(\{x_1, \dots, x_N\})$  subject to the constraint  $g(\{x_1, \dots, x_N\}) = 0$  is obtained by solving the  $N + 1$  equations :

$$\frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} = 0 \quad ; \quad i \in \{1, \dots, N\}, \quad (0.25)$$

$$g(\{x_1, \dots, x_N\}) = 0 \quad . \quad (0.26)$$

## 0.4 Method of steepest descents

In general it is extremely difficult to integrate functions which occur in exponentials. The method of steepest descents is a rather useful tool which is frequently used to approximate certain types of integrals. In our analysis of statistical mechanics we often wish to compute integrals of the following type:

$$\int_{-\infty}^{+\infty} dx f(x) = \int_{-\infty}^{+\infty} dx e^{\ln f(x)} = \int_{-\infty}^{+\infty} dx e^{g(x)}. \quad (0.27)$$

Since  $g(x)$  appears in an exponent, the integral will be dominated by the parts of the domain close to the maximum of  $g(x)$  which will contribute exponentially more to the integral than those elsewhere. The maximum of  $g(x)$  occurs when

$$\left. \frac{dg(x)}{dx} \right|_{x_0} = 0 \quad (0.28)$$

and if

$$\left. \frac{d^2 g(x)}{dx^2} \right|_{x_0} < 0. \quad (0.29)$$

Expanding  $g(x)$  in a Taylor series around the maximum  $x_0$  gives

$$g(x) = g(x_0) + \frac{1}{2} (x - x_0)^2 g''(x_0) + \dots \quad (0.30)$$

such that one can use Gaussian integration to approximate

$$\begin{aligned} \int_{-\infty}^{+\infty} dx f(x) &= \int_{-\infty}^{+\infty} dx e^{g(x)} \simeq e^{g(x_0)} \int_{-\infty}^{+\infty} dx \exp \left[ -\frac{1}{2} (x - x_0)^2 |g''(x_0)| \right] \\ &= e^{g(x_0)} \sqrt{\frac{2\pi}{|g''(x_0)|}} \quad , \end{aligned} \quad (0.31)$$

where we have used the fact that  $g''(x_0) < 0$ .

# 1 Thermodynamics ( or what makes thermodynamics different from classical mechanics?)

Thermodynamics is the description of macroscopic systems where **heat** and **temperature** play a role. **Equilibrium thermodynamics** is the study of such **macrosystems** in **thermal equilibrium**.

This course will help us to understand the meaning of all those ‘technical’ terms in the sentences above. Let us start by describing the meaning of some of them.

**Definition:** A **thermodynamic system** is a bounded macroscopic quantity of matter (macroscopic here means that the size of the system is comparable to our human scale). The **boundary surface** need not have a fixed shape or volume and separates the system from its **environment** (e.g. when a gas expands against a moving piston, the volume of the gas enclosed by the boundary surface changes).

We will typically wish to study scenarios in which there is a flow of matter and energy into or out of the system<sup>2</sup>.

- An **isolated** system cannot exchange energy or particles (matter) with the environment.
- A **closed** system can exchange energy but not particles with its environment (such as heat conduction or doing work).
- An **open** system can exchange both matter (particles) and energy with the environment

We can characterise such a thermodynamic system by **state variables**. These are macroscopic quantities like volume,  $V$ , pressure,  $p$ , number of ‘particles’,  $N$ , temperature,  $T$  and chemical potential,  $\mu$  (in contrast to microscopic quantities like the masses, positions and momenta of the particles). It is natural to call them *variables* since they can be *varied* in experiment.

Some state variables are proportional to the size of the system and are called **extensive** variables (e.g. volume,  $V$  and number of particles,  $N$ ). Other state variables are independent of size (the amount of matter) in the system and are called **intensive** variables (e.g. pressure,  $p$ , chemical potential,  $\mu$  and temperature,  $T$ )

Many of the state variables form intensive-extensive **conjugate pairs** such that their product has dimensions of energy. The intensive member of each pair is a force-like variable while the extensive member of the pair has the properties of a displacement. Typically we refer to a **generalized force** and a **generalized displacement**.

Under a *constant* generalized force,  $\mathbf{f}$  the work done in moving a generalized displacement  $\mathbf{x}$  is

$$\Delta W = \mathbf{f} \cdot \mathbf{x}$$

---

<sup>2</sup>In this course we will be concerned with non-relativistic systems for which we can make a clear distinction between matter and energy.

| System              | Generalized Force, $\mathbf{f}$ | Generalized Displacement, $\mathbf{x}$ |
|---------------------|---------------------------------|--|
| Gas                 | pressure, $-p$                  | volume, $V$                            |
| Elastic filament    | tension, $\tau$                 | length, $L$                            |
| Magnet              | flux density, $\mathbf{B}$      | dipole moment $\mathbf{m}$             |
| Liquid film         | surface tension, $\gamma$       | area, $A$                              |
| Dielectric material | electric field, $\mathbf{E}$ ,  | dipole moment, $\mathbf{p}$            |
| Chemical species    | chemical potential, $\mu$ ,     | number of molecules, $N$               |

## 1.1 What is Equilibrium?

A system is in **thermodynamic equilibrium** when its properties do not change appreciably during the time when it is under observation. Clearly this definition is subjective as it depends on *how long* we observe it (e.g panes of glass in Medieval churches).

A fundamental assumption of thermodynamics (or the thermodynamical approach) is the idea that **any isolated system** after **sufficient time** settles into a **stationary state**, called the **equilibrium state** in which **all the state variables** are **constant**. The system is then said to be in **thermodynamic equilibrium**. This is a very reasonable assumption (which has been obtained empirically from experiment) and we have plenty of experience (over the course of human history!) which makes such an assumption quite intuitively obvious.

Hence we can characterise **equilibrium states** by **functions of the state variables**.

The phrase *sufficient time* implies the existence of a typical timescale (the **relaxation time**). When observing a system we would have to wait longer than the relaxation time to be sure that the system is in equilibrium.

Given the subjective nature of our definition, a suitably (scientifically) pragmatic definition of a system in equilibrium would then be when **all** the state variables being used to describe the system were not changing. If each state variable has its own specific relaxation time, then this would then require us to wait for timescales longer than all the relaxation times of the particular state variables being used to describe the system<sup>3</sup>. That is, we would have to wait for longer than the longest relaxation time in the system (the relaxation time of the ‘slowest’ state variable).

In addition, bringing two bodies/systems into contact (restricting ourselves for now to the case where they *do no work on each other*) will in general lead to changes in the values of their state variables. When all the state variables are no longer changing, then the 2 systems are said to be in **thermal equilibrium**.

## 1.2 What are Heat and Temperature ? - The Zeroth and First Laws

To continue our study of macroscopic systems we need to understand what is meant by **heat** and **temperature** and the difference (if any) between them. We all have an intuitive idea of the difference between a hotter and a colder body but we will need to

---

<sup>3</sup>This also leaves open the possibility for us to study systems where certain macroscopic quantities are at equilibrium while others are not - the field of non-equilibrium thermodynamics, an exciting area of active research with many fundamental open questions, unfortunately outside the scope of this course

develop a more precise idea of these notions <sup>4</sup>

### 1.2.1 The Zeroth Law of Thermodynamics

We can ask ourselves the following question. Is it possible for 2 bodies to **not be** in contact and be in thermal equilibrium with each other? Upon reflection, the answer is yes. However one would need a 3rd body (such as a thermometer) to verify if that was so. The zeroth law encapsulates this fact <sup>5</sup>.

The zeroth law states :

If two macroscopic bodies are in thermal equilibrium with a third, then they are also in thermal equilibrium with each other.

Let us consider the implications of this.

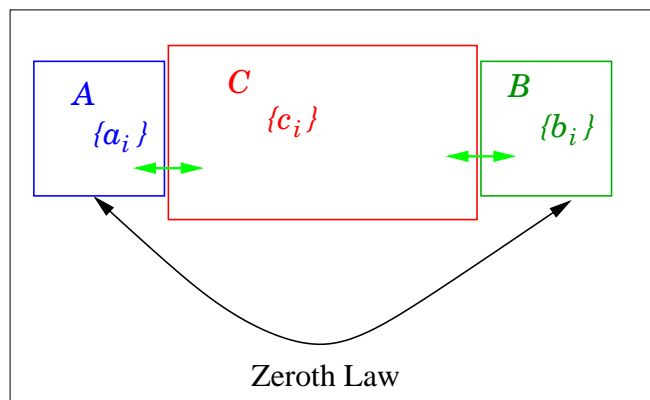


Figure 1: The Zeroth Law: 2 systems  $A, B$  which are both separately in thermal equilibrium with a third system  $C$  must be in equilibrium with each other.

Consider two systems  $A, B$  which are both separately in thermal equilibrium with a third system  $C$ . (Note that we consider thermal equilibrium so we do not consider the case where they do work on each other). The zeroth law then implies that  $A$  and  $B$  must be in equilibrium with each other.

The equilibrium states of  $A, B, C$  are described by their state variables  $(a_1, a_2, \dots)$ ,  $(b_1, b_2, \dots)$ ,  $(c_1, c_2, \dots)$  respectively.

If  $A$  is in equilibrium with  $C$ , then this equilibrium can be expressed as a constraint on the variables of  $A$  and  $C$ , i.e. if a change is made to one of the state variables of  $A$  then some change must occur in the variables of  $C$  to keep the two systems in equilibrium. This constraint can be expressed as :

$$f_{AC}(a_1, a_2, \dots; c_1, c_2, \dots) = 0 \Rightarrow c_1 = F_{AC}(a_1, a_2, \dots; c_2, c_3, \dots) \quad (1.1)$$

---

<sup>4</sup>There is an interesting discussion on the history of the development of the concept of heat on the BBC Radio 4 programme "In our time" with Melvyn Bragg on December 4, 2008. The episode was called "Heat: A History - from fire to thermodynamics" You can download a podcast on [www.bbc.co.uk/radio4/history/inourtime/inourtime\\_science.shtml](http://www.bbc.co.uk/radio4/history/inourtime/inourtime_science.shtml).

<sup>5</sup>This law was only recognized as a logical necessity for thermometry only in 1931 long after the other laws were stated

The equilibrium of  $B$  and  $C$  implies a similar constraint:

$$f_{BC}(b_1, b_2, \dots; c_1, c_2, \dots) = 0 \Rightarrow c_1 = F_{BC}(b_1, b_2, \dots; c_2, c_3, \dots) \quad (1.2)$$

Both these equations imply that

$$F_{AC}(a_1, a_2, \dots; c_2, c_3, \dots) = F_{BC}(b_1, b_2, \dots; c_2, c_3, \dots) \quad (1.3)$$

which implies that

$$a_1 = f_{ABC}(b_1, b_2, \dots; a_2, a_3, \dots; c_2, c_3, \dots) \quad (1.4)$$

However the Zeroth Law requires that the following is true :

$$f_{AB}(a_1, a_2, \dots; b_1, b_2, \dots) = 0 \Rightarrow a_1 = F_{AB}(b_1, b_2, \dots; a_2, a_3, \dots) , \quad (1.5)$$

*independent* of the set of variables  $\{c_i\}$ .

Now both equations (1.4) and (1.5) give an expression for  $a_1$  in terms of the other state variables so they must be the same. However equation (1.5) is independent of the set of variables  $\{c_i\}$  which therefore implies that all the  $\{c_i\}$  variables **must** drop out of equation (1.4). As a result they must **also** drop out of equation (1.3) <sup>6</sup> which means :

$$F_{AC}(a_1, a_2, \dots) = F_{BC}(b_1, b_2, \dots) , \quad (1.6)$$

Hence, there must exist a function only of the state variables of  $A$  which is equal to a function of only the state variables of  $B$  when  $A$  and  $B$  are in thermal equilibrium with each other .

$$\Theta_A(a_1, a_2, \dots) = \Theta_B(b_1, b_2, \dots) \equiv \Theta , \quad (1.7)$$

The thermal equilibrium of the 2 systems  $(i, j)$  is characterised by functions  $\Theta_i, \Theta_j$  of their respective state variables being the same for both of them <sup>7</sup>.

This function,  $\Theta$  is a **state function** since it is a function of state-variables. Note that it is therefore also constant at equilibrium which means that it could also be used as a *variable* in an experiment.

This function specifies an **equation of state** linking the state variables. From what we have developed to this point, we do not have a clear idea of what  $\Theta$  is - only that it exists. As we go on we will see that what we intuitively know as the **temperature** is a perfect candidate for this function/variable.

This is often called the definition of the **empirical temperature**.

Clearly all macroscopic equilibrium states of  $A$  at a ‘temperature’  $T$  fall on the curve  $\Theta_A(a_1, a_2, \dots) = T$  which is called an **isotherm**.

---

<sup>6</sup>Note that this must be true for any system and hence for arbitrary functions  $F_{AB}, F_{BC}$ .

<sup>7</sup>Note the functional form of  $\Theta_i$  may be different for the two systems e.g. two different metals

**Example** We can get an intuitive idea of the Zeroth law by considering its application to the following: a set of systems each being a gas enclosed in a cylinder by a frictionless piston (state variables:  $p_i, V_i$ ). We have 3 systems,  $i = 1, 2, 3$  and let us choose system-3 as our reference state with fixed  $p_3, V_3$ .

Figure 2: Isotherms in the p-V plane.

If we require that system-1 is in equilibrium with system-3, this constraint leaves only one independent variable e.g. for each  $p_1$  there is a unique  $V_1$  which is a function of  $p_3, V_3, p_1$  and we can plot on the p-V plane all values of  $p_1, V_1$  that are in equilibrium with 3. These points will all lie on a curve on the p-V plane which has no overhangs. Let us call it an isotherm and label it by  $\Theta_1$ . Similarly if we require that system-2 is in equilibrium with system-3, we will get a similar curve which we can also label, this time by  $\Theta_2$ . The zeroth law requires that systems 1 and 2 are also in thermal equilibrium with each other which means that all points lying on curve  $\Theta_1$  must be in equilibrium with all points lying on  $\Theta_2$  (and with each other). They must have a property in common which allows this to be so - the temperature.

We could have similarly used a similar argument to the zeroth law for mechanical and chemical equilibrium to show existence of a ‘generalized force’ and chemical potential as *intensive* state variables that are identical in any pair of bodies in mechanical or chemical equilibrium respectively.

### 1.2.2 The ideal gas and thermodynamic temperature

Therefore if we find a system where we can keep all variables except *one* constant, with no mechanical work done, then changes in that *one* variable can be used to define changes in  $\Theta$ . The value of that one variable (and hence  $\Theta$ ) **must** be the same between two copies of such a system which are in equilibrium. We can then use this system unambiguously to determine how ‘hot’ any other body is.

To establish a temperature scale then all we need is a material with a property that varies (almost) linearly with temperature (like the height of mercury in a thin tube), choose 2 easily reproducible reference points (e.g. freezing point/boiling point of water at 1 atm) and define a linear scale between them.

Dilute gases form a particularly insightful class of thermometers.

1. It is found empirically that at low constant pressures all gases expand more or less linearly with temperature,  $T = aV + b \Rightarrow V = (T - b)/a$ .
2. At low densities, it is also found that the constant of proportionality (between  $V$  and  $T$ ) for all gases with the same numbers of particles is *identical* irrespective of the chemical structure of the particles making up the gas.
3. If one extrapolates to low temperatures, there seems to be a unique temperature,  $T_0$  at which all gases would seem to have zero volume (in reality they would condense before that). This temperature is called **absolute zero**.

So if we define a temperature scale such that at absolute zero, the temperature is zero, then all dilute gases satisfy (to some approximation) the equation

$$pV = RT \tag{1.8}$$

where  $T = 0$  corresponds to **absolute zero**.

This is called the **ideal gas** law. In reality there are always deviations from this law for real gases, the deviations becoming smaller at lower pressures. However, such idealised situations are the always excellent (dare one say, ideal) starting points for understanding the behaviour of physical systems (recall Newton's 1st law).

Since we only need one temperature to fix this scale, we can use the triple point of water (where water-ice-steam co-exist in equilibrium) which occurs at a unique pressure and temperature and define it to be 273.16 units. Then it turns out that boiling point and freezing point of water at 1 atm are 373.15 and 273.15 units respectively. So absolute zero is  $-273.15$  °C. The units are called Kelvin (K) after Lord Kelvin (W. Thomson) who suggested fixing the scale this way <sup>8</sup>.

A temperature scale that is defined such that at  $T = 0$  at absolute zero, has temperature as always **positive**, i.e.  $T \geq 0$  and is called a **thermodynamic** temperature scale.

The zeroth law clearly allows many different temperature scales, however we will see later that the second law implies this unique (up to a scale factor) temperature scale.

Equation (1.8) is an example of an **equation of state**, which is a relationship between state variables  $F(p, V, T) = 0$  that is valid at equilibrium.

We note that since  $V$  is extensive then so is  $R$ .

In SI units  $R = k_B N$  where  $k_B = 1.38 \times 10^{-23}$  J/K is called Boltzmann's constant and  $N$  is the number of particles in the gas.

For 1 mole of a gas, the number of particles is given by  $N_A = 6.02 \times 10^{23}$ , called **Avogadro's number**. 1 mole of a substance is defined as the amount of a substance that contains as many elementary entities as there are carbon atoms in 0.012kg of Carbon-12 ( $^{12}C$ ). The mass of 1 mole of any substance is given by its molecular mass  $\times N_A \times 1$  Da <sup>9</sup> (remember the molecular mass of hydrogen,  $^1H$  is 1, oxygen,  $^{16}O$  is 16 and water  $H_2O$  is 18). Hence one atom of  $^{12}C$  weighs 12 atomic mass units ( $12u$ ), one atom of  $^1H$  weighs  $1u$ , one atom of  $^{16}O$  weighs  $16u$  and so on.

<sup>8</sup>W. Thomson, *Phil. Trans. Roy. Soc. A*, **144**, 350, (1854)

<sup>9</sup>1 atomic mass unit or Dalton (Da) often denoted as  $u$  is defined as 1/12 of the mass of 1 atom of  $^{12}C = 1.66 \times 10^{-27}$ kg.

### 1.2.3 Thermodynamic Reversibility

A process is **reversible** if and only if its direction can be reversed by an **infinitesimal** change in the conditions. Note that the process is NOT reversible if it requires a **finite** change to reverse it.

A reversible process is a process that, **after** it has taken place, can be reversed - returning the system to its initial state where both the system and environment recover exactly their properties at the beginning of the process.

Thermodynamic reversibility requires that

1. the process is **quasistatic** and
2. there is no **hysteresis**

**Quasistatic processes** These are processes which are carried out so slowly that every state through which the system goes can be considered an equilibrium state (i.e. the process should be infinitely slow, however for pragmatic operational reasons it actually must only be much slower than the slowest relaxing process). Fast changes lead to different parts of the system not being in equilibrium with each other. For example, consider moving a piston of a gas cylinder to compress a gas inside it. If this is done fast, sound/shock waves will be set up creating regions of different pressure/temperature. Clearly this is irreversible since moving the piston out does not reverse the sound waves. Quasistatic processes can be described using continuous curves in **state space**. Non-quasistatic processes cannot be so illustrated. Since all points on the process are at equilibrium if we stop at any point in the process and the system does not change.

**Hysteresis** For a system with hysteresis, when a quasistatic process is reversed, the system does not trace its previous path. A common source of hysteresis is friction.

Figure 3: A quasistatic process showing hysteresis.