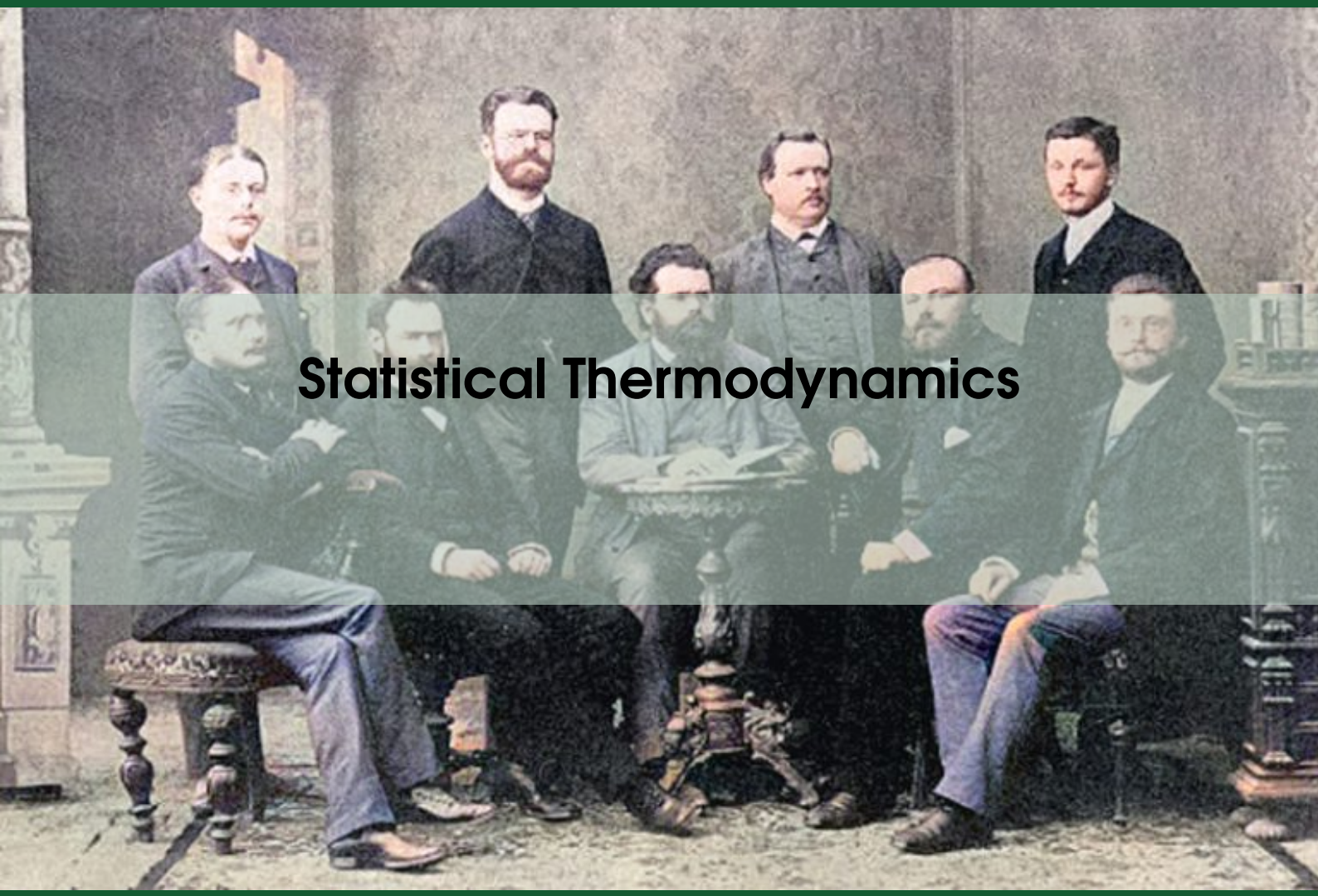
Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

Colección CBI

Lecture Notes



Statistical Thermodynamics

Orest Pizio
Víctor M. Trejos

Statistical Thermodynamics

Orest Pizio
Víctor M. Trejos

Lecture Notes in Statistical Thermodynamics

Statistical Thermodynamics

Primera edición: 2024

Av. Ferrocarril San Rafael Atlixco, Núm. 186, Col. Leyes de Reforma 1 A Sección, Alcaldía Iztapalapa, C.P. 09310, Ciudad de México

ISBN Colección: 978-607-28-2107-1
ISBN Volumen: 978-607-28-3687-7

Contents

Preface

Very many textbooks on statistical thermodynamics are available. Each of them is marked by the points of view and pedagogical insights of the authors into practically the same set of basic topics and elements. In general terms, we have been motivated by different factors. Three decades passed since publication of the textbook by Leopoldo Garcia Colin "Statistical thermodynamics" in 1995 [1]. For years, it served as a valuable instrument for teaching the subject at the Autonomous Metropolitan University of Mexico. That book apparently has been written for students of physics at the graduate level. Our intention, however, is to reconsider and extend the subject by combining and explaining various topics differently. Present lecture notes are designed as pedagogical tools either to teach or to study statistical thermodynamics. Amount of material covered by these lectures is adjusted according to the requirements of the programs of postgraduate studies in chemistry at the National Autonomous University of Mexico and at Autonomous Metropolitan University of Mexico. Within these programs, the statistical thermodynamics is the obligatory subject that is taught during one semester or during two trimesters, respectively. Thus, the proposed text contains twenty five lectures (called chapters in the book format). Unfortunately, the programs of study do not envisage additional hours to solve problems or exercises. Therefore, in several lectures we offer homeworks for students to develop their practical skills. On the other hand, we elaborated supplementary materials mostly containing mathematical elements that are used in the principal text. As noted above, this set of lecture notes is designed for students of different levels of non-physics faculties, like chemists and chemical engineers. The lectures can be useful for those studying biophysics and biochemistry or students of mathematics to learn some methodological aspects and applications of this part of physics. The lecturer and students of physical chemistry would need the course as one of principal theoretical instruments and as practical tools. In contrast, the students of other areas of chemistry can find it helpful for solving particular problems that involve inter-disciplinary aspects. One can consider these lectures as an introduction permitting to take later a more sophisticated, advanced courses concerned with the topics of quantum mechanics, of the statistical theory of liquids and solutions or with computer simulation methods, during e.g. Ph.D. studies. For the moment, basis amount of knowledge is provided making possible to proceed to advanced problems of statistical thermodynamics, if necessary.

Thermodynamics is a part of physics that deals with concepts commonly known as heat, work, and temperature, and their relation to energy and properties of matter. This

branch of science permeates and applies to a wide variety of phenomena in nature and engineering. Formally it is a part of physics, but applies to several problems of chemistry, biology and other sciences. In order to make things clear, first note that statistical physics is a branch of theoretical physics that involves mechanics and mathematical statistics. This part of physics considers matter as composed from a large number of particles (atoms or molecules, for example) and consequently requires statistical concepts to describe the entire system in question. Statistical physics can be called or defined as kinetic theory of matter as well, because it involves motion of individual particles that leads to collective phenomena like transfer of heat or of matter within the macroscopic system. Formally, the statistical physics can be divided into statistical thermodynamics that describes physical systems solely at thermodynamic equilibrium and statistical kinetics focused in the theory of processes occurring in such physical systems.

The text is divided into 25 chapters, beginning with the first two chapters, which introduce the Lagrangian and Hamiltonian formulations of classical mechanics. The next three chapters provide an introduction to ensemble averages, the Maxwell distribution of velocities, and related applications. Chapter 6 discusses intermolecular interactions and describes reduction of the pair interaction potential into the one-particle external field. Chapters 7-9 offer an introduction to quantum mechanics, and the solution of basic quantum problems. Chapters 10-12 introduce ensemble theory, covering the canonical, grand canonical, and isobaric-isothermal ensembles, along with discussions of the classical mechanical limit. Chapters 13-14 focus on Boltzmann statistics, with detailed discussions of the partition function, thermodynamic properties of monoatomic and polyatomic gases, and rotational and vibrational partition functions. Chapters 15-16 delve into the statistical thermodynamics of paramagnetic materials, the one-dimensional Ising model, and the molecular field method for ferromagnetism and electrolyte solutions. In Chapter 17, the Boltzmann distribution is used to describe the Langmuir adsorption isotherms of fluids on surfaces. Chapters 18-19 explain the Fermi-Dirac and Bose-Einstein statistics and blackbody radiation. The final chapters 20-24 include discussions of the Tonks-Takahashi one-dimensional model fluid, the partition function of non-ideal classical systems of particles, and concepts of liquid theory used in phase equilibrium calculations. The last chapter 25 deals with the concepts of chemical association. In addition, eight appendices are provided to support the mathematical elements of the course.

In general terms the content of the course in majority of its elements follows from the book by H.T. Davis "*Statistical Mechanics of phases, interfaces, and thin films*" [2]. However, due to temporal restriction of the semester the content covers roughly only the first five chapters of [2], substantially redone according to our understanding of the subject. The quantum mechanics part, has been adapted using the books on quantum mechanics by A.S. Davydov [3], and I.O. Vakarchuk [4]. On the other hand, chapter 16 of the present work, as well as all the elements of the theory of liquids have been elaborated from the textbooks published in Ukrainian and Russian, principally in the times of the former USSR. We used some elements of the books by V.B. Kobyliansky [5], and Yukhnovsky and Holovko [6]. These resources are unavailable outside their countries of origin, but enrich these lecture notes in our opinion. Selected problems offered to the reader are adapted from McQuarrie [7, 8]. Finally, we have also intended to incorporate methodological developments and some materials from the works of Leopoldo García Colin [1] and Fernando del Río [9].

Orest Pizio
Víctor M. Trejos
Ciudad de México
November, 2024

1. Introduction

We will study the subject known as statistical thermodynamics. In order to explain its contents, let us begin from some general introductory comments. Let us say first, that statistical mechanics represents a branch of theoretical physics that studies macroscopic systems from a microscopic, or molecular, point of view. Deepness of the microscopic insight is determined by a problem one need to solve. Obviously, fundamental approach is microscopic by construction. Because the systems under study in everyday life are macroscopic, i.e. they contain an enormous number of particles, we need to design and apply some sort of statistical approach.

So what we wish to do is to start from the concepts of particles and inter-particle interactions and to construct a theory for a macroscopic system. In other words, the principal goal of statistical mechanics is to provide understanding (or say interpretation) on one hand, and make predictions on the other hand, both concerned with macroscopic phenomena that emerge from the properties of individual molecules making up the system. What is taken into account is the intra-molecular structure and inter-particle interactions. To summarize: one intends to obtain macroscopic properties and their changes dependent on the state of the system from the microscopic insight. Many of the macroscopic outputs can be obtained from the experiments directly or indirectly. Thus, it is desirable to perform comparisons of theoretical predictions with experimental data.

There are several kinds of properties we would like to focus in. Those are thermodynamic, dynamic, electric, magnetic, optical and others. Examples may include: the equation of state, heat capacities and compressibilities, speed of sound, refractive index, dielectric constant and related relaxation phenomena, thermal and electric conductivity, self-diffusion, viscosity. All that can be obtained. The route is: first try to understand with the aim to predict next (for example materials with novel or desirable or controllable properties).

The techniques (or methods) of statistical mechanics have been used (or applicable) while studying variety of physical problems or phenomena and different kinds of systems. They can be applied to gases and compressed gases, liquids, solutions, electrolyte solutions, polymers, membranes, metals and alloys. A few examples of phenomena are: adsorption on solids and wetting, phase transitions, diffusion and transport. Each of these words actually represents a huge area of research.

Formally, in fact dependent on what is studied, statistical mechanics can be divided into two parts: one dealing with macroscopic systems in equilibrium, let us call it the sta-

tistical thermodynamics (it can be considered as a bridge between molecular physics and common thermodynamics) and the second is denominated as non-equilibrium statistical mechanics. Common thermodynamics yields (provides) formal mathematical relations between experimentally accessible (measurable) properties of macroscopic systems in equilibrium. These properties are frequently hidden in general thermodynamic relations.

One frequently used example is:

$$C_p - C_v = \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right]\left(\frac{\partial V}{\partial T}\right)_p, \tag{1.1}$$

where C_p and C_v are the heat capacities at constant pressure and constant volume, respectively. Other variables are P, V, and T (denoting pressure, volume and absolute temperature) and the internal energy U. The properties in the left hand side (l.h.s.) of this equation are experimentally measurable for certain substance. The variables P, V and T are coupled between themselves by the equation of state. The P and T, or V and T variables can be chosen as the conditions of observation. The meaning of the derivatives in the right hand side (r.h.s.) of Eq. (1.1) is quite obscure.

The Eq. (1.1) is an exact relation. However, thermodynamics on its own does not provide information about the magnitudes of the properties involved, unless a model for the substance is assumed. Thermodynamics does not rely upon existence of atoms or molecules in contrast to statistical thermodynamics. Transformation of the Eq. (1.1) into the form that involves solely experimentally measurable quantities is given in the Appendix 26.1.

Statistical thermodynamics is easy to apply to systems in which the inter-particle interactions and their effect on the relevant properties can be neglected, e.g. dilute gases. Still statistics is involved in such treatment. On the other extreme, statistical thermodynamics is necessary, but becomes more difficult methodologically, if one attempts to deal with "strongly correlated" systems, for example liquids, electrolyte solutions, etc.

If we consider a system of many particles from mechanical point of view, it is necessary to solve equations of motion describing time evolution of the positions (say configurations) of all molecules. Time is the principal variable in such construction. How to solve such problem is the first question here.

State of the system of N particles, or mechanical state, let call it a micro-state, at certain moment of time is given by specifying all the coordinates ($3N$ in Cartesian space) and $3N$ components of velocities (or momenta). Description of the system means or requires to determine the micro-state as well as its evolution in time. Note, that the energy of the system is determined by the coordinates and velocities, the potential and kinetic energy, respectively. Here, it is worth to recall from methodological point of view that in quantum mechanics, the coordinates and momenta cannot be described precisely simultaneously. Finally, it is evident that one does not have a device to measure or follow changes of the micro-state of many-particle system. In contrast, in order to define a macro-state, or a state of the system for say coarse-grained, thermodynamic description one needs to define only a few parameters (external and/or internal, for example: if pressure is external - volume of the system results, if external electric or magnetic fields are given then polarization or magnetization as internal characteristics follow).

Prior to attempt progress along this line of reasoning, let us make some comments concerning the system we will deal with. For simplicity, we restrict our attention to a pure (one-component) homogeneous system of N particles in a volume V. It is assumed that the system is in thermal equilibrium. In other words, the system has been equilibrated due to the contact with a heat bath (huge reservoir) at temperature T. From the point of view of thermodynamics, one can characterize the system at such conditions by giving its extensive properties like total energy (commonly denoted as U or E), the volume, V, and the number of particles, N (or the total mass M). The intensive variables are: $\rho = M/V$ (or $\rho = N/V$ - the number density), $\tilde{U} = U/M$ (or $\tilde{U} = U/N$). The intensive variables are

constants throughout one-phase, uniform system. One more issue: the system should be macroscopic to possess thermodynamics, i.e. the thermodynamic limit is valid, $N \to \infty$, $V \to \infty$, such that $N/V = const$.

So, in order to develop a microscopic description of the system it is necessary to specify completely the state of N particles using classical or quantum mechanics. What are the options? Start from the problem of one particle, however. The equation of motion is the second-order differential equation. Historically first, Newton formulation has been developed. If we have N particles, the equations of motion for each of them can be written as,

$$m_i \left(\frac{d^2 \mathbf{r}_i}{dt^2} \right) = \mathbf{F}_i(r_1, r_2, ..., r_N), \quad i = 1, ..., N \tag{1.2}$$

where m_i is the mass of the particle and $\mathbf{r}_i$ is the vector describing the coordinates of a particle i in the laboratory frame. On the other hand, $\mathbf{F}_i$ is the force that acts on the particle i due to all other particle and possibly due to any external field if present. The force is assumed conservative, i.e. it depends only on the coordinate but not on time. So, what do we have? Very many, or better say too many equations to solve. Besides, each of equations requires two boundary conditions. One, formally useful observation is that, if one defines momentum of a particle, $\mathbf{p}$, rather than velocity, by using, $\mathbf{p} = m\mathbf{v} = md\mathbf{r}/dt$, then the second-order equations convert into the first-order differential equations for this auxiliary quantity, such that,

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i(r_1, r_2, ..., r_N). \tag{1.3}$$

It is assumed above that mass is independent on time (say no relativistic effects, here it is worth to consult D. Landau, Statistical Physics [10]). Another comment is the following: we are speaking about the motion of a small particle in the same way as about the motion of a macroscopic body, applying the same mechanics.

D. McQuarrie in his book on statistical thermodynamics [7] used three simple examples for illustration or application of Newton's mechanics. The first of them is the following exercise.

Exercise 1.1 Describe the motion of a body with the mass, m, shot vertically upward from the Earth level (at z_0) with an initial velocity, v_0, in a gravitational field. Thus, the boundary conditions for the initial coordinate and velocity are given a priori. The motion is one-dimensional. If we recall that the potential energy of a body in the gravitational scalar field is $U = mgz$ and the conservative force is related to the potential energy as, $\mathbf{F} = -\partial U(\mathbf{r})/\partial \mathbf{r}$, then the solution of the problem is given as a second-order polynomial,

$$z(t) = -\frac{1}{2} g t^2 + v_o t + z_0. \tag{1.4}$$

■

Another, frequently used example, necessary in our future developments in different parts of these chapters, is the harmonic oscillator.

Exercise 1.2 Again, it is an example concerned with solely a one-dimensional motion. Consider a spring, rigidly attached to a solid body at one end and having a mass, m, at the opposite end. At equilibrium, the length of the spring is x_0. If a force, not very big by magnitude, is applied either to compress or extend the spring along the axis x, normal to the solid surface, the force and change of spring length are related by the Hooke's law, $F = -k(x - x_0)$, where k denotes the elasticity constant, it is the property of the material of which the spring is done. Note, that the elastic potential energy is

$U = k(x - x_0)^2/2$. Introduce, a rescaled coordinate, $\xi = x - x_0$. We need to solve the following differential equation,

$$m\frac{d^2\xi}{dt^2} = -k\xi, \tag{1.5}$$

that again is easy to solve and find solution in the form,

$$\xi(t) = Asin\omega t + Bcos\omega t, \tag{1.6}$$

where $\omega^2 = k/m$, and ω denotes the characteristic, or natural, frequency of the oscillator. Two unknown constants in the solution, A and B are determined by the initial conditions. Alternatively, the solution can be written in equivalent form as $Csin(\omega t + \phi)$, where C is the amplitude and ϕ is the phase of oscillations.

■

The following, last example is more elaborated.

Exercise 1.3 It concerns a two-dimensional motion of a body under the influence of Coulomb-type attraction, i.e. the potential is $\sim 1/r$, to a fixed center. Thus, the force is $\mathbf{F} = -\alpha\mathbf{r}/r^3$, where α is a parameter. The equation of motion, $m\ddot{\mathbf{r}} = \mathbf{F}$, in the Cartesian coordinates ($\mathbf{r} = \mathbf{i}x + \mathbf{j}y$, and $\mathbf{F} = \mathbf{i}F_x + \mathbf{j}F_y$) are,

$$m\ddot{x} = F_x = -\frac{\alpha x}{(x^2 + y^2)^{3/2}}, \tag{1.7}$$

and

$$m\ddot{y} = F_y = -\frac{\alpha y}{(x^2 + y^2)^{3/2}}, \tag{1.8}$$

where, F_x and F_y are the projections of the force on the coordinate axes, x and y. These are two coupled differential equations difficult to solve by hands. However, there exists a circular symmetry of the problem that can be used to make things simpler. Let us make transformation to polar coordinates and use them rather than the Cartesian ones: $x = rcos\theta$, $y = r\sin\theta$. After calculating derivatives ($\dot{x}$ and $\ddot{y}$) and substituting them into Eqs. (1.7) and (1.8), we obtain:

$$[m(\ddot{r} - r\dot{\theta}^2) + \alpha/r^2]cos\theta - m(r\ddot{\theta} + 2\dot{r}\dot{\theta})sin\theta = 0, \tag{1.9}$$

$$[m(\ddot{r} - r\dot{\theta}^2) + \alpha/r^2]sin\theta + m(r\ddot{\theta} + 2\dot{r}\dot{\theta})cos\theta = 0. \tag{1.10}$$

Multiplying the first of these resulting equations by $\cos\theta$ and the second by $\sin\theta$ and next adding them we observe that,

$$[m(\ddot{r} - r\dot{\theta}^2) + \alpha/r^2] = 0. \tag{1.11}$$

Then,

$$m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = 0, \tag{1.12}$$

as well. This latter Eq. (1.12) can be rewritten in the compact form as,

$$\frac{1}{r}\frac{d}{dt}(mr^2\dot{\theta}) = 0. \tag{1.13}$$

Therefore, $mr^2\dot{\theta} = const$, and is the integral of motion. This quantity is called as the angular momentum l, $l = mr^2\dot{\theta}$. It is conserved during motion under the influence of the so-called central force, i.e. the force that is directed along the vector $\mathbf{r}$, defined in the statement of the problem. In consequence, the Newton's equation of motion in Cartesian coordinates, $m\ddot{r} = F$, keeps its form but the force is given as $F = -\alpha/r^2 + l^2/mr^3$, as it follows from Eq. (1.11). It will become the Newton's equation, if we interpret l^2/mr^3 as an additional, centrifugal force. Not always, interpretation of the terms for forces appearing upon transformation of coordinates is as straightforward as in the present exercise. To conclude, the form $m\ddot{\xi} = F_{\xi}$ is useful in Cartesian coordinates only, unless we are able or prepared to define some additional forces.

■

A more convenient formulations of mechanics, or say of the equations of motion, that do not change upon transformation of the coordinates, are the Lagrangian and Hamilton constructions. Those are considered in the next chapter.

Exercise 1.4 We have used the following concepts: energy and forces. Forces come from the potential energy. Newtons's equation of motion is the relation between potential energy and kinetic energy, because on one side we have force and on another velocity. Now the question is: how to illustrate the relation between work and kinetic energy. Let us do such an exercise for one-dimensional motion. Start from the Newton's equation: $F = ma = m\frac{d}{dt}(\frac{dx}{dt})$. Assume that the volume is $V = Sx$ ($dV = Sdx$), where S denotes area. By definition, the pressure is $P = F/S$ and the work then is written as PdV. Thus,

$$PdV = \left(\frac{F}{S}\right) Sdx = (ma)dx = m\frac{d}{dt}\left(\frac{dx}{dt}\right)dx = m\left(\frac{dv}{dt}\right)dx. \tag{1.14}$$

where $v = dx/dt$. Then, the energy is:

$$E = \int PdV = m\int\left(\frac{dv}{dt}\right)\frac{dx}{dt}dt = m\int vdv = mv^2/2, \tag{1.15}$$

showing what we have searched for.

■

Exercise 1.5 A) Convert a point (-1,-1) given in Cartesian coordinates into polar coordinates (r,θ);

B) Convert the equation $r = -8cos\theta$ into Cartesian coordinates and show that it is a circle of radius 4 and centered at (-4,0);

C) The height of a body shot vertically upward is given by the equation $h(t) = 40t - 32t^2$. How high it will go?

D) Differentiate the function,

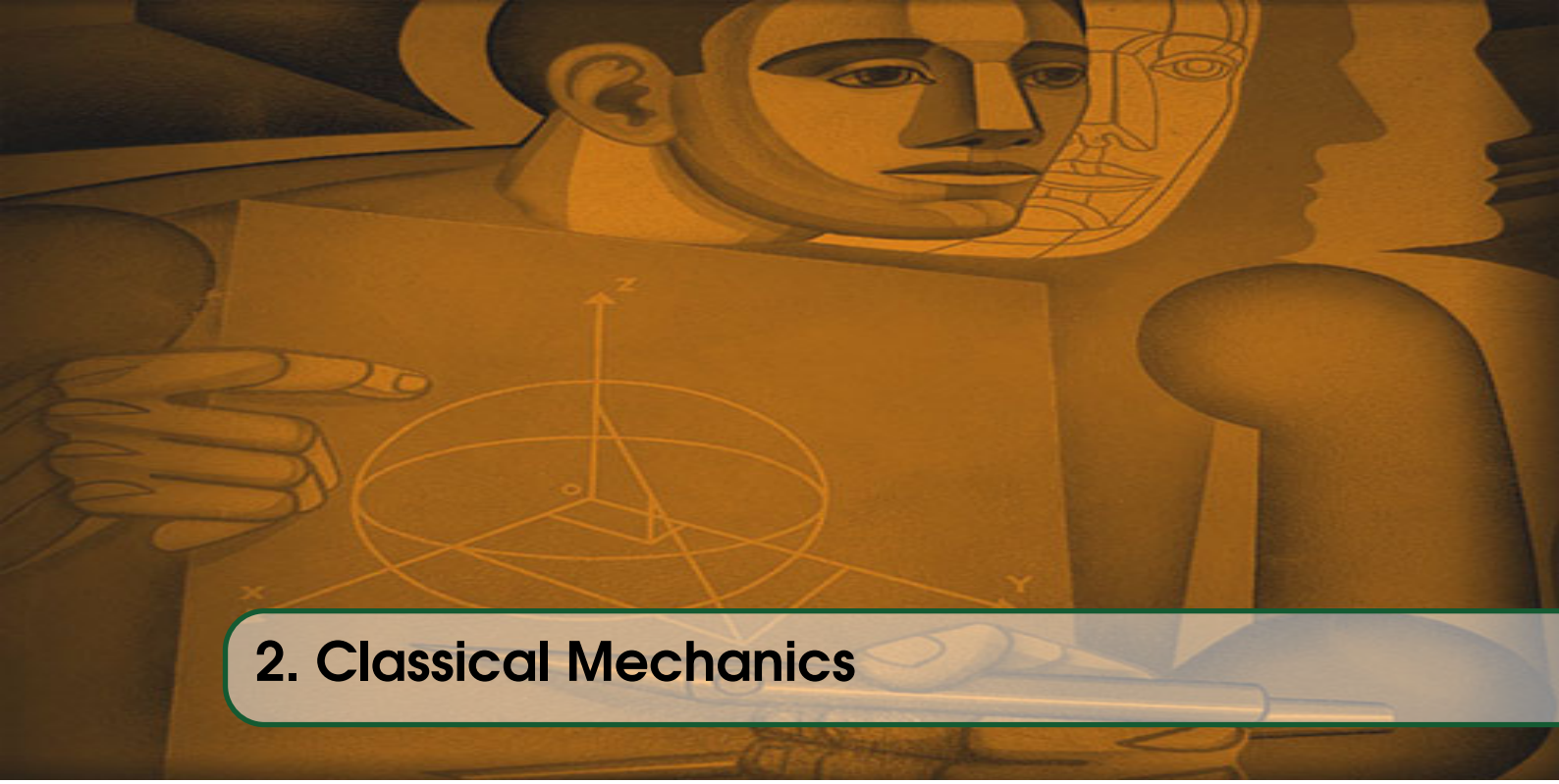
$$y(x) = x^2e^{-x}cosx. \tag{1.16}$$

E) Find dy/dx if,

$$y(x) = \exp(-(x^2 + a^2)^{1/2}), \tag{1.17}$$

where a is a constant.

■



2. Classical Mechanics

2.1 Lagrangian formulation

We proceed now to the formulation of mechanics by using the concept of Lagrangian function and of Hamiltonian. Begin with Lagrangian formulation. The principal idea is in deriving equations of motion without resorting to the concept of forces. Recall that the kinetic energy, K, of a single particle in three-dimensional space in Cartesian coordinates is,

$$K(\dot{x},\dot{y},\dot{z}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2). \tag{2.1}$$

It is assumed that a particle is under influence of conservative forces, this means that its potential energy, U, is a function of coordinates only, i.e. $U = U(x,y,z)$. The potential energy is a result of the influence of other particles of the system or due to the external field acting on the entire system, e.g. gravitational field. Let us introduce or define the function, called as the Lagrangian,

$$L = K - U = K(\dot{x},\dot{y},\dot{z}) - U(x,y,z). \tag{2.2}$$

Then,

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial K}{\partial \dot{x}} = m\dot{x}, \tag{2.3}$$

and

$$\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x}. \tag{2.4}$$

Therefore, the equation of motion, if written in the Newton's form, $m\ddot{x} = -\partial U/\partial x$, can be expressed as,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}. \tag{2.5}$$

Similar equations can be written for y and z coordinates. Hence, for a single particle we have three equations of motion, involving three components of the Cartesian coordinate

and of velocity. If one assumes a set of arbitrary coordinates, q_i, for N-particle system, call them the generalized coordinates, then the Lagrangian equations of motion are,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}, \tag{2.6}$$

with $i = 1,...,3N$. The Lagrangian is a function of generalized velocities and generalized coordinates, i.e. $L = L(\dot{q}_1,...,\dot{q}_{3N}; q_1,...,q_{3N})$. This is a very important statement.

Let us show now, that the equations of motion in the Lagrangian form are invariant upon the transformation of coordinates, in contrast to Newton's equations. Explore this issue by considering a single particle in three-dimensional space for simplicity. Imagine that we made transformation of "old" coordinates, x,y,z, into a set of arbitrary new ones, q_1,q_2,q_3, by using the equations: $x = x(q_1,q_2,q_3)$, $y = y(q_1,q_2,q_3)$ and $z = z(q_1,q_2,q_3)$. Then, for the r.h.s. of Eq. (2.6) we have,

$$\frac{\partial L}{\partial q_1} = \frac{\partial L}{\partial x}\left(\frac{\partial x}{\partial q_1}\right) + \frac{\partial L}{\partial y}\left(\frac{\partial y}{\partial q_1}\right) + \frac{\partial L}{\partial z}\left(\frac{\partial z}{\partial q_1}\right) + \frac{\partial L}{\partial \dot{x}}\left(\frac{\partial \dot{x}}{\partial q_1}\right) + \frac{\partial L}{\partial \dot{y}}\left(\frac{\partial \dot{y}}{\partial q_1}\right) + \frac{\partial L}{\partial \dot{z}}\left(\frac{\partial \dot{z}}{\partial q_1}\right). \tag{2.7}$$

The l.h.s. of Eq. (2.6) can be manipulated as follows,

$$\frac{\partial L}{\partial \dot{q}_1} = \frac{\partial L}{\partial x}\left(\frac{\partial x}{\partial \dot{q}_1}\right) + \frac{\partial L}{\partial y}\left(\frac{\partial y}{\partial \dot{q}_1}\right) + \frac{\partial L}{\partial z}\left(\frac{\partial z}{\partial \dot{q}_1}\right) + \frac{\partial L}{\partial \dot{x}}\left(\frac{\partial \dot{x}}{\partial \dot{q}_1}\right) + \frac{\partial L}{\partial \dot{y}}\left(\frac{\partial \dot{y}}{\partial \dot{q}_1}\right) + \frac{\partial L}{\partial \dot{z}}\left(\frac{\partial \dot{z}}{\partial \dot{q}_1}\right). \tag{2.8}$$

However, the coordinates x,y,z do not depend on new velocities, $\dot{q}_1,\dot{q}_2,\dot{q}_3$. Thus, the first three terms in Eq. (2.8) are zero. Moreover, it can be observed that, $\partial \dot{x}/\partial \dot{q}_1 = \partial x/\partial q_1$. Then,

$$\frac{\partial L}{\partial \dot{q}_1} = \left(\frac{\partial L}{\partial \dot{x}}\right)\left(\frac{\partial x}{\partial q_1}\right) + \left(\frac{\partial L}{\partial \dot{y}}\right)\left(\frac{\partial y}{\partial q_1}\right) + \left(\frac{\partial L}{\partial \dot{z}}\right)\left(\frac{\partial z}{\partial q_1}\right). \tag{2.9}$$

With this result, we can transform the l.h.s. of Eq. (2.6) further, namely by taking necessary time derivative,

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) = \left(\frac{\partial x}{\partial q_1}\right)\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\right] + \left(\frac{\partial y}{\partial q_1}\right)\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right)\right] + \left(\frac{\partial z}{\partial q_1}\right)\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{z}}\right)\right] + \\ \frac{\partial L}{\partial \dot{x}}\left(\frac{\partial \dot{x}}{\partial q_1}\right) + \frac{\partial L}{\partial \dot{y}}\left(\frac{\partial \dot{y}}{\partial q_1}\right) + \frac{\partial L}{\partial \dot{z}}\left(\frac{\partial \dot{z}}{\partial q_1}\right).\end{aligned} \tag{2.10}$$

The multipliers in the square brackets in the first three terms of this equation should by substituted according to the Lagrange equations of motion in old coordinates, Eq. (2.5). Then, by comparison of transformed in such a way Eq. (2.10) and Eq. (2.7), we obtain the desired result,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) = \frac{\partial L}{\partial q_1}, \tag{2.11}$$

Similar procedure should be performed for the variables q_2 and q_3.

Let us comment now what we have. After introducing by definition the Lagrange function, L, we have derived the equations of motion that involve this function. One can think that this is just a reformulation of Newton's equations of motion. However, now we do not need the concept of forces. Rather, the kinetic and potential energies are involved. The Lagrange equations of motion are the second-order differential equations. To solve them a set of boundary conditions is needed. Namely, the initial velocities and initial coordinates (as in Newton's formulation). One can attempt to develop or construct statistical mechanics of many particles by using Lagrangian function.

As an illustration of the application of a new form of equations of motion, one can redo the last exercise from the chapter 1 (concerned with the circular motion of a particle around attractive center). Recall that in polar coordinates, $x = rcos\theta$ and $y = rsin\theta$. Then,

$$\dot{x} = \dot{r}cos\theta - rsin\theta \cdot \dot{\theta}; \quad \dot{y} = \dot{r}sin\theta + rcos\theta \cdot \dot{\theta}, \tag{2.12}$$

and the Lagrangian function for this problem is in the form,

$$L = K - U = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - U = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{\alpha}{r}. \tag{2.13}$$

Then the equation of motion involving polar coordinates are:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{d}{dt}(mr^2\dot{\theta}) = \frac{\partial L}{\partial \theta} = 0, \tag{2.14}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = \frac{d}{dt}(m\dot{r}) = \frac{\partial L}{\partial r} = mr\dot{\theta}^2 - \frac{\alpha}{r^2}. \tag{2.15}$$

The Eq. (2.14) yields conservation of the angular momentum, while Eq. (2.15) is the desired equation of motion,

$$m\ddot{r} = mr\dot{\theta}^2 - \frac{\alpha}{r^2}. \tag{2.16}$$

All the trick was just to express the Lagrangian function in terms of appropriate coordinates. Still, to solve the equations of motion one needs to apply certain boundary conditions.

2.2 Hamiltonian formulation

The Lagrange equations of motion are the second-order differential equations and require the boundary conditions for the velocities and coordinates, say $6N$ constants for the system of N particles. It is seducing to put the following question: is it possible to apply more equations but of lower order than the Lagrange's equations? This is precisely how the Hamiltonian formulation of mechanics is done.

Assuming that the Lagrange's function, L is given, we would like to eliminate velocities and introduce the generalized momenta. By definition, the generalized momenta are,

$$p_i = (\partial L/\partial \dot{q}_i), \quad i = 1,2,..,3N \tag{2.17}$$

It is said, that p_i is conjugate to q_i, e.g. $p_x = m\dot{x} = (\partial L/\partial \dot{x})$. Now, let us define the Hamiltonian by,

$$H(p,q) = \sum_{i=1}^{3N} p_i\dot{q}_i - L(\dot{q};q), \quad i = 1,2,....,3N \tag{2.18}$$

where $\dot{q}$ in L should be eliminated using the definition in Eq. (2.17). To proceed, we recall that the kinetic energy in general can be written as,

$$K = \sum_{i=1}^{3N} a_i(q_1,q_2,\cdots,q_{3N})\dot{q}_i^2, \tag{2.19}$$

where the parameters a_i usually are just constants and do not depend on the coordinates explicitly. With this form of kinetic energy we have,

$$p_i = (\partial L/\partial \dot{q}_i) = (\partial K/\partial \dot{q}_i) = 2a_i\dot{q}_i. \tag{2.20}$$

Then,

$$H(p,q) = \sum_{i=1}^{3N} p_i\dot{q}_i - L = 2\sum_{i=1}^{3N} a_i\dot{q}_i^2 - L = K + U. \tag{2.21}$$

So, we see that the Hamiltonian just is the total energy.

From the first part of this chapter, we should note that the Lagrange's function does not depend on time explicitly. Having this in mind, we would like to show that $dH/dt = 0$, i.e. that the Hamiltonian does not depend on time as well. According to the definition given by Eq. (2.18) for the Hamiltonian, we see that,

$$dH = \sum_i \dot{q}_i dp_i + \sum_i p_i d\dot{q}_i - \sum_i (\partial L/\partial \dot{q}_i) d\dot{q}_i - \sum_i (\partial L/\partial q_i) dq_i. \tag{2.22}$$

Now, the definition of the momenta according to Eq. (2.17) can be used to get,

$$dH = \sum_i \dot{q}_i dp_i - \sum_i (\partial L/\partial q_i) dq_i \tag{2.23}$$

Next, we apply the Lagrangian equations of motion given by Eq. (2.6) to arrive at the result,

$$dH = \sum_i \dot{q}_i dp_i - \sum_i \dot{p}_i dq_i. \tag{2.24}$$

On the other hand, we know that the Hamiltonian is the function of momenta and coordinates, $H = H(p,q)$. Thus,

$$dH = \sum_i (\partial H/\partial p_i) dp_i + \sum_i (\partial H/\partial q_i) dq_i. \tag{2.25}$$

Therefore, it follows from the comparison of Eqs. (2.23) and (2.25) that,

$$(\partial H/\partial p_i) = \dot{q}_i; \quad (\partial H/\partial q_i) = -\dot{p}_i. \tag{2.26}$$

We see that the evolution of momenta and of coordinates in time is determined by the derivatives of the Hamiltonian. Two equations in Eq. (2.26) are the equations of motion in the Hamiltonian form. For the system of N particles, there are $6N$ first-order differential equations that require $6N$ boundary conditions.

Consequently,

$$dH/dt = \partial H/\partial t + (\partial H/\partial p)(dp/dt) + (\partial H/\partial q)(dq/dt) = 0, \tag{2.27}$$

because of equations of motion given by Eq. (2.26). Thus, the Hamiltonian is the integral of motion. The Hamilton formulation of mechanics is conceptually most convenient for the development of statistical mechanics.

Exercise 2.1 Offered as a homework: Necessary mathematical elements for a few next steps are: 1) transformation from Cartesian coordinates to cylindrical and spherical coordinates, 2) Jacobian of transformation between different sets of coordinates, 3) gradient in spherical coordinates. ■

2.3 From mechanics to thermodynamics

The simplest thermodynamic system serving as a cornerstone for further developments is an ideal gas. For sufficiently dilute gases it has been found experimentally that the ratio $P/[(N/V)T]$ equals to a universal constant k_B for various gases composed of monoatomic molecules. Here, N/V is the number density of particles of a gas (it should be quite small), and k_B is the Boltzmann constant. In other words, the equation of state simply is,

$$PV = Nk_BT. \tag{2.28}$$

On the other hand, it is convenient to refer to experimental observations to claim that the energy of a dilute monoatomic gas denoted as E, is,

$$E = \frac{3}{2} Nk_BT. \tag{2.29}$$

Concerning the notion of ideal gas, we refer to a macroscopic system of N point particles with mass m in a macroscopic volume V. The principal issue is that the number density of particles is such that the average distance between any pair of particles is very big. It is an imaginary construction that physically is realized in very dilute gases. In such physical system, one can mention that the distance between particles should be much larger than the range of interparticle interaction. We restrict our attention to one-component systems unless specified.

If the kinetic energy of a single particle i is $mv_i^2/2$, then the total energy of the system at a given instant of time is,

$$E = \frac{1}{2}\sum_{i=1}^{N} mv_i^2. \tag{2.30}$$

It is assumed that any external influence on the system is absent but the walls of the container determining the volume V. Moreover, to make things clear, we would like to note another important assumption. We assume that **average** energy of the system is constant (i.e. it is a thermodynamic system in equilibrium) but the instantaneous values of energy can change with time either due to collisions of particles with the walls or not frequent but possible inter-particle collisions. So, there exists certain mechanism of the redistribution of energy in the system. Referring to any measurement performed on real, even very dilute system, the observer obtains an average value of energy that can be formally written as follows,

$$\overline{E}_\tau = \frac{1}{\tau}\int_0^\tau E(t)dt. \tag{2.31}$$

Here, the bar denotes the time average and τ is the time interval of the measurement. As a postulate, it is assumed that a thermodynamic quantity equals to the long time average of the corresponding microscopic quantity. So, in the case of a dilute, monoatomic gas the energy of the system (in thermodynamic sense) is,

$$E = \frac{1}{2}\sum_{i=1}^{N} m\overline{v_i^2} = \frac{1}{2}Nm\overline{v^2}, \tag{2.32}$$

where $m\overline{v^2}/2$ is the average kinetic energy of any particle of the gas. The sum has been formally performed because all the particles are equivalent. The notion of time average in the equation above expresses everything what we in fact do not know about the behavior of the system or more specifically about the initial distribution of the coordinates of particles and of the time evolution of their velocities.

Comparing this result above with the experimentally grounded observation expressed by Eq. (2.29), we obtain,

$$\frac{1}{2}m\overline{v^2} = \frac{3}{2}k_B T. \tag{2.33}$$

To conclude, it appears that the average kinetic energy of a particle in dilute gas at thermal equilibrium is proportional to the absolute temperature. Using this relation, the link between the kinetic energy and the equation of state can be established. Namely,

$$P = \frac{1}{3V}\sum_{i=1}^{N} m\overline{v_i^2} = \frac{1}{V}\sum_{i=1}^{N} m\overline{v_{ix}^2} = Nk_B T/V, \tag{2.34}$$

because the system in question is isotropic in absence of any external field, and thus,

$$\overline{v_i^2} = \overline{v_{ix}^2 + v_{iy}^2 + v_{iz}^2}, \quad \overline{v_{ix}^2} = \overline{v_{iy}^2} = \overline{v_{iz}^2}. \tag{2.35}$$

How to perform the time average is an open question at the moment. Actually, it is impossible to do, because we would need to measure in practice velocity of a particle along sufficiently large interval of time for many-particle system or in theory to solve equations of motion.

Time averaging is a consequence of microscopic approach to describe thermodynamic systems and to deal with time evolution of micro-states. Hence, alternative approaches should be constructed. Intuitively, one would think to take advantage of the concept of probability.

3. Maxwell distribution of velocities

3.1 Towards ensemble averages

In the previous chapter we have shown that the equation of state (EOS) of an ideal gas is the result of time averaging of the kinetic energy of the system, see Eq. (2.34) of chapter 2. The question left open is: what does it mean? and/or how to perform such average? The time averaging is required because of application of microscopic approach to a thermodynamic system. This averaging reflects "smoothing" of the description of the time evolution of microscopic states for a given system subjected to thermodynamic constraints.

In order to explain the calculation of an integral over time in Eq. (2.31), we would like to give the following reasoning. Assume, that a system is characterized by a random variable, K (it is a kinetic energy in our case), which takes a sequence of values, $K_1, K_2, ...$ (discreteness is assumed for simplicity) as a function of time. The time of observation, τ, should be large enough to obtain a reasonable result after this kind of measurement because at short times, certain relaxation processes can occur resulting in a drift of an average. However, more important is that the values for K fluctuate even after the system is well equilibrated. So, we imagine formally that the values $K_1, K_2, ...$ can be observed during time intervals $t_1, t_2, ...$, respectively. With this discretized version, the time averaging leads to the result,

$$\overline{K} = \sum_i K_i \frac{t_i}{\tau}. \tag{3.1}$$

If τ is quite large, and the number of intervals, i, is very big as well, one can intuitively interpret t_i/τ as the probability P_i for occurring the value K_i. The overline over K in the (l.h.s.) of Eq. (3.1), denotes this kind of average, i.e. the time average over microstates that change upon time. Recall, the the entire physical system in question is considered or encountered under certain conditions, e.g. N, V, T. The system is considered in equilibrium. The number of thermodynamic parameters or macroscopic constraints describing the state of a system is just three for one-component case. Concerning the constraints, we are not restricted to N, V, T. Namely, the N, V, E, or N, P, T, or μ, V, T (μ is the chemical potential) macroscopic constraints can be applied.

Let us now, think in another way and perform gedankenexperiment (imaginary or thought experiment). Namely, imagine a set of the systems (the same as the original)

all subject to the same macroscopic constraints or equivalently under the same thermodynamic conditions. For example, take a set of NVT equal systems, all equilibrated by contact with the same thermal bath. All these copies may refer to certain NVT original system and are called replicas. Apparently, all these replicas are equivalent for external observer because they represent copies of a system of our interest under the same constraints.

The microscopic state of all particles (as determined by the coordinates and momenta of particles) in any of these replicas is certainly different and unknown or inaccessible, however. The number of replicas of a system in the imaginary construction, M, is very big. It should be so big, such that there are $\nu_1, \nu_2, \nu_3, ...$ copies in which a random variable, K, takes values, $K_1, K_2, K_3, ...$ The number of copies, M, should be very big to include All possible values of K that are observed upon evolution of a single original system in time. A set of replicas is called an ensemble. It is important to mention, that the imaginary experiment for producing replicas does not involve time variable, say we pick up all replicas at arbitrary instant of time.

In order to put these two points of view together, time evolution and replicated single system under the same constraints, we note that for a given random variable, K,

$$\begin{aligned} &\frac{t_i}{\tau} \to P_i; \\ &\frac{\nu_i}{M} \to P_i, \end{aligned} \tag{3.2}$$

for $\tau \to \infty$ and $M \to \infty$, both yielding the probability to get the value K_i. Hence, the long-time average and ensemble average for certain characteristics of thermodynamic system can be expressed in terms of probability calculated differently. The ensemble average is usually denoted by $< ... >$, so we can use the equation above to get,

$$\overline{K} =< K >= \sum_i K_i P_i. \tag{3.3}$$

This is the principal postulate of the ensemble theory. It is also known as ergodic hypothesis. The ensemble theory always yields the results in agreement with thermodynamics, this can be considered as an implicit proof of the validity of such theory. At the moment, we have introduced a new construction, but how to perform ensemble average is not specified.

Therefore as an illustration, let us make the following exercise. The EOS of an ideal gas involves thermodynamic parameters T, and P. Also it contains the number density of a gas, $n = N/V$. Evidently, the temperature has statistical meaning. Intuitively, the pressure, P, has statistical meaning as well, as it is a mechanical property related to the velocities or the kinetic energy of particles in a dilute gas.

The pressure is the force per area, the force is the ingredient of Newton's mechanics via equation of motion. In close similarity to the Exercise 1.4 of chapter 1, we write $F/S \to m\frac{d}{dt}(\frac{dx}{dt})/S \to \frac{\Delta p}{\Delta t}/S$, where the momentum, p, $p = mv$. It is the relation between the moment, Δp, given to a unit surface during time interval Δt. In such setup, it is assumed that a particle moves in the direction of x axis and can strike or collide with the surface in the yz plane.

The momentum change upon elastic collision of a single particle with the surface is $2p_x$, where p_x is the projection of the momentum p on x axis. We have noted above that for a many-particle system we need to calculate the ensemble average to obtain a property of interest. Assume that there exists certain, unknown for the moment, distribution of velocities or equivalently of momenta of particles.

In contrast to comments given above for the case of discrete variable, we would like to deal with continuous variable and apply machinery similar to the last part of previous chapter 2.

Let us assume that there exists the probability density distribution of particles according to the value of their momenta and denote it by $w(p_x)$. Then, $dw(p_x) = w(p_x)dp_x$ is the probability that x-component of momentum of a particle is in the interval $(p_x, p_x + dp_x)$.

The number of particles with momenta in this interval is $ndw(p_x)$ where $n = N/V$ is the number density of a gas as already mentioned above. Denote the time interval of our observation as τ. Not all the particles from the system collide with the surface, but only those that are sufficiently close to it. In other words, the particles reach the surface in time τ only, if they are at most at a distance $(p_x/m)\tau$. Thus, the total change of momenta per time interval yielding the pressure would be,

$$\left[(2p_x)(ndw(p_x))\left(\frac{p_x}{m}\right)\tau\right]/\tau. \tag{3.4}$$

The pressure comes out from the particles with arbitrary values of momenta, directed to the surface,

$$P = \frac{2n}{m}\int_0^\infty (p_x)^2 dw(p_x). \tag{3.5}$$

The key issue that we have not studied so far, is to have expression for $dw(p_x)$ or equivalently to know how the momenta of particles are distributed probabilistically. Let us take the probability density distribution of momenta, $w(p_x)$, from, say molecular physics, and use the following expression,

$$w(p_x) = \frac{1}{(2\pi m k_B T)^{1/2}} \mathrm{e}^{-p_x^2/2mk_BT}. \tag{3.6}$$

It is worth mentioning that the distribution function has exponential form. The kinetic energy enters with negative sign, actually the ratio of kinetic energy and thermal energy is in the exponent. Then, the pressure can be calculated straightforwardly,

$$P = \frac{2n}{m}\frac{1}{(2\pi m k_B T)^{1/2}}\int_0^\infty (p_x)^2 \mathrm{e}^{-p_x^2/2mk_BT} dp_x = nk_BT. \tag{3.7}$$

So, what is the trick? The answer is, take the probability distribution and express desirable thermodynamic property using the distribution.

A brief description of the probability theory is given in Appendix 26.2.

Exercise 3.1 Offered as a homework: Perform the integration of the type, $\int dx x^2 \exp[-cx^2]$. Prior to that calculate the Gaussian integral $\int dx \exp[-cx^2]$ by using transformation from Cartesian coordinates to polar coordinates. ■

3.2 Maxwell distribution

Let us continue with theoretical developments for dilute gas using the ideal gas model. Assume that the gas is under N, V and T conditions. Consider a particle belonging to a system of an ensemble (chosen at random) and assume that the particle has velocity components in the small interval $(v_x, v_x + \Delta v_x)$, $(v_y, v_y + \Delta v_y)$ and $(v_z, v_z + \Delta v_z)$, around the point v_x, v_y, v_z. The volume element describing the interval is $\Delta v_x \Delta v_y \Delta v_z = \Delta^3 v$.

Our intention is to obtain the probability distribution $p(\mathbf{v}, \Delta^3 v)$. It is determined through the probability density $\phi(\mathbf{v})$,

$$\phi(\mathbf{v}) = \lim_{\Delta v_x, \Delta v_y, \Delta v_z \to 0} \frac{p(\mathbf{v}, \Delta^3 v)}{\Delta v_x \Delta v_y \Delta v_z}. \tag{3.8}$$

If the limit above is performed, we have

$$p(\mathbf{v};d^3v) = \phi(\mathbf{v})d^3v. \tag{3.9}$$

The probability density should satisfy the normalization condition over the entire space of velocities,

$$\int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_z \phi(\mathbf{v}) = 1. \tag{3.10}$$

In order to proceed, we need now to derive the function $\phi(\mathbf{v})$. According to the original ideas of Maxwell it has been assumed that:

1. The function $\phi(\mathbf{v})$ depends on velocity solely through the kinetic energy, i.e. it depends on v^2 rather than on $\mathbf{v}$ itself,

$$\phi(\mathbf{v}) = \phi(mv^2/2), \tag{3.11}$$

This assumption expresses that the system is isotropic, i.e. there is no preferential direction.

2. The components of the kinetic energy in Cartesian coordinates contribute equally and independently into the function $\phi(\mathbf{v})$, i.e.

$$\phi(\mathbf{v}) = \gamma(mv_x^2/2)\gamma(mv_y^2/2)\gamma(mv_z^2/2), \tag{3.12}$$

where $\gamma(x)$ is an unknown function and $mv^2/2 = m(v_x^2 + v_y^2 + v_z^2)/2$.

3. Final assumption is that the function $\gamma(x)$ is differentiable. In some sense it means that the acceleration is well defined as it is the derivative of velocity.

With these assumptions, the form of the density probability function can be obtained rather straightforwardly. To begin with, consider the differentiable functions f and g with the following property (in the spirit of assumption 2 above),

$$f(x+y) = g(x)g(y), \tag{3.13}$$

for any values of arguments x and y. Take derivative of this equation by the argument x

$$f'(x+y) = g'(x)g(y), \tag{3.14}$$

and by y,

$$f'(x+y) = g(x)g'(y). \tag{3.15}$$

Then,

$$g'(x)g(y) = g(x)g'(y), \quad or \quad \frac{g'(x)}{g(x)} = \frac{g'(y)}{g(y)}. \tag{3.16}$$

The arguments x and y are independent, in consequence the equation above implies that

$$\frac{g'(x)}{g(x)} = -\beta, \tag{3.17}$$

where β is arbitrary constant. The solution of the first-order differential equation is,

$$g(x) = \alpha e^{-\beta x}. \tag{3.18}$$

Here α is the integration constant, unknown for the moment. Repeating similar development for the function dependent on three arguments, $f(x+y+z) = g(x)g(y)g(z)$, one can easily obtain similar result as above.

Consequently, returning to the assumption number 2 of Maxwell, we can write down the density distribution function $\phi(\mathbf{v})$ as,

$$\phi(\mathbf{v}) = \phi(mv^2/2) = \alpha^3 e^{-\beta m v^2/2}. \tag{3.19}$$

The unknown constant of integration α can be determined from the normalization condition for the probability density distribution,

$$\int d^3v\phi(\mathbf{v}) = \alpha^3 \int\int\int dv_x dv_y dv_z e^{-\beta m(v_x^2+v_y^2+v_z^2)/2} = 1. \tag{3.20}$$

Calculation of three equal by form integrals, over v_x, v_y and v_z, is necessary to perform (exercise requested above: obtain the result of integration $\int dxdy \exp[-c(x^2+y^2)]$ by using polar coordinates). As the result, we get the constant of integration,

$$\alpha = (\beta m/2\pi)^{1/2}. \tag{3.21}$$

Still, the parameter β is not determined. To get it, recall Eq. (2.33), $m\overline{v^2} = 3k_BT$. Then, the result is,:

$$\beta = (k_BT)^{-1}. \tag{3.22}$$

Therefore, the final form of the Maxwell distribution is,

$$\phi(\mathbf{v}) = \left(\frac{m}{2\pi k_BT}\right)^{3/2} \exp\left(-\frac{mv^2}{2k_BT}\right). \tag{3.23}$$

Our result refers to an isothermal system, dilute, or say ideal gas, composed of particles with the mass m. The dispersion of the distribution (characteristic width) is determined by the statistical parameter T and by mass of particles. Because, N particles are independent, we conclude that for the entire system, the probable number of particles having velocities in the interval $\mathbf{v}$ and $\mathbf{v} + d^3v$ is $N\phi(\mathbf{v})dv_xdv_ydv_z$.

In close similarity with the development above, it is worth noting that any kind of average of arbitrary function of velocities, $h(\mathbf{v})$ can be calculated using the probability density distribution as,

$$< h(\mathbf{v}) >= \int d^3v h(\mathbf{v})\phi(\mathbf{v}). \tag{3.24}$$

This expression can be used in several aspects. Namely, for example we need to describe the probability that the x component of velocity of a particle lies in the interval, $v_x, v_x + dv_x$, regardless of the values of velocity along y and z directions. Then,

$$\phi_x(v_x)dv_x = dv_x \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dv_y dv_z \phi(\mathbf{v}) = \left(\frac{m}{2\pi k_BT}\right)^{1/2} \exp\left(-\frac{mv_x^2}{2k_BT}\right) dv_x. \tag{3.25}$$

Next, with the result above, it can be easily shown that,

$$< v_x >= \int_{-\infty}^{\infty} dv_x v_x \phi_x(v_x) = 0. \tag{3.26}$$

Similarly, $< v_y >=< v_z >= 0$.

Exercise 3.2 A particle has the point dipole moment μ. It is free to rotate with respect to (w.r.t.) the center of mass. Assume the center of mass (COM) frame in spherical coordinates, with angles θ and φ; $0 < \theta < \pi$, $0 < \varphi < 2\pi$. The projection of the dipole moment on the z-axis is μ_z, $\mu_z = \mu cos\ \theta$. The probability that a dipole is oriented in the solid angle $sin\ \theta\ d\theta d\varphi$ is,

$$dp = \omega sin\ \theta\ d\theta d\varphi, \tag{3.27}$$

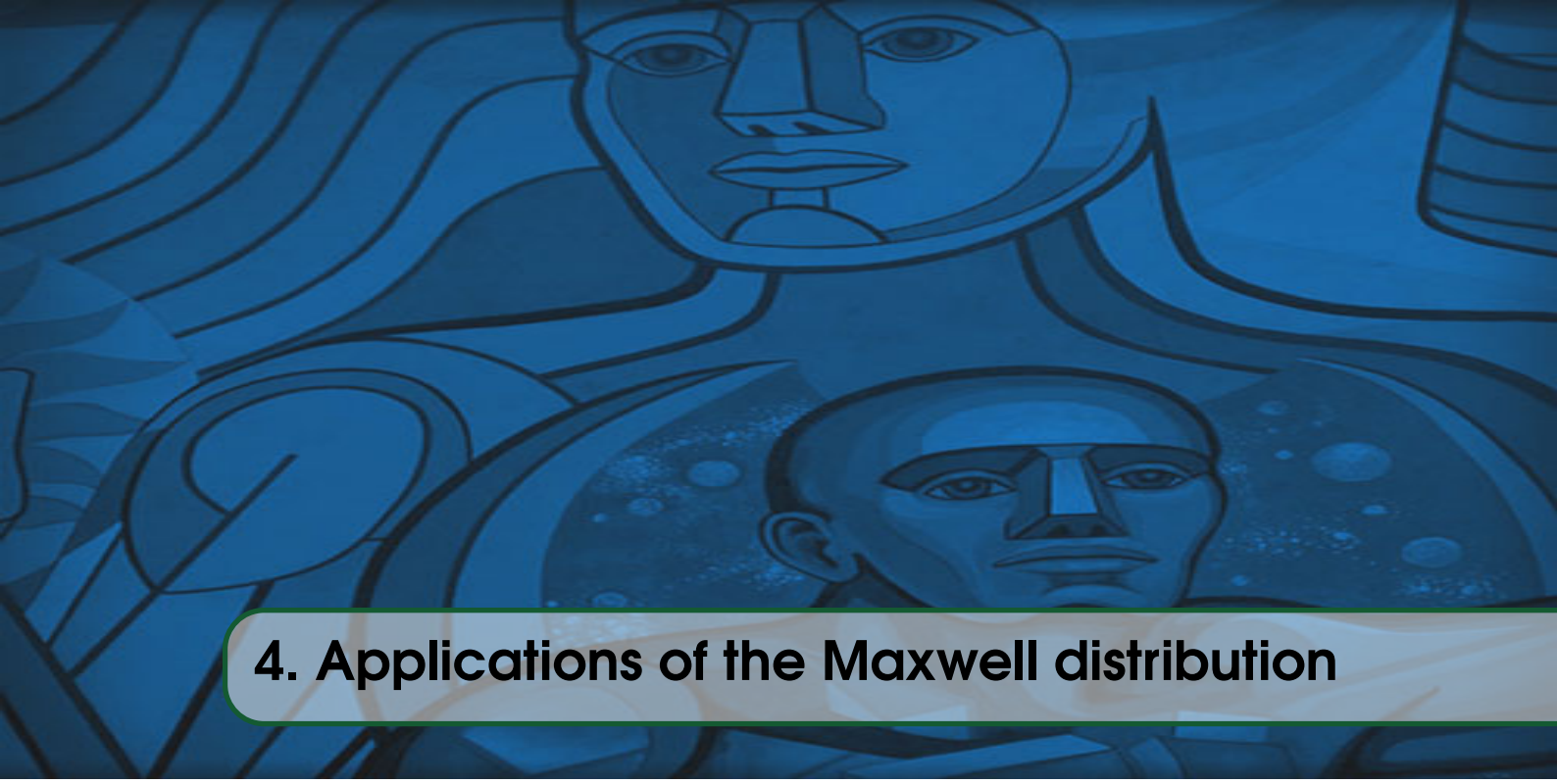
where the probability density, ω, is $\omega = 1/4\pi$. Calculate the mean $< \mu_z >$ and the mean square $< \mu_z^2 >$ of the projection of the dipole moment. ■

Exercise 3.3 With the given probability density distribution,

$$\omega = \frac{\mu E \exp(-\mu E cos\ \theta / k_B T)}{2\pi k_B T \left[\exp(\mu E / k_B T)\right] - \exp\left(-\mu E / k_B T\right)}. \tag{3.28}$$

Calculate the mean $< \mu_z >$ and the mean square $< \mu_z^2 >$ of the projection of a dipole moment under the influence of an external electric field, E, directed along z-axis.

■



4. Applications of the Maxwell distribution

4.1 Some applications of the Maxwell distribution

In the Appendix 26.2, Eq. (26.26), we have used the result for the continuous probability density distribution of a real random variable in the form,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{4.1}$$

It is called as normal or Gaussian distribution. However, the general form of this function is,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right], \tag{4.2}$$

where the parameter μ is the mean (or expectation value) of the distribution, and the parameter σ is denominated as its standard deviation. The variance of the distribution is σ^2. Thus, the Eq. (4.1) is the simplest case of a normal distribution with $\mu = 0$ and $\sigma = 1$ called as the standard normal distribution.

For the sake of comparison, recall that the one-dimensional Maxwell distribution for the velocity component along, for example x axis, is,

$$\phi(v_x) = \left(\frac{m}{2\pi k_B T}\right)^{1/2} \exp\left(-\frac{mv_x^2}{2k_B T}\right). \tag{4.3}$$

Thus, $\mu = 0$ and the standard deviation $\sigma = (k_B T/m)^{1/2}$.

Mathematical note: 1) definition of the cumulative distribution function (CDF) of the standard normal distribution is like we used in Appendix 26.2, Eq. (26.34),

$$F_{\xi}(x) = \int_{-\infty}^{x} f_{\xi}(t)dt. \tag{4.4}$$

Then, for normal distribution,

$$F_{\xi}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt. \tag{4.5}$$

2) It is intrinsically related to the frequently used special function, called the error function (denoted as erf(x)) that involves integration of the normal distribution of mean $\mu = 0$ and variance $\sigma^2 = 1/2$,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \tag{4.6}$$

This function has several applications in statistical mechanics.

End Mathematical note.

In physics, we use the term called root mean square (rms) width or rms deviation, it is equivalent to the variance of the density distribution. In the case of the x-component of velocity, it is,

$$v_x^{rms} = (< (v_x - < v_x >)^2 >)^{1/2}. \tag{4.7}$$

Note, from the right to left: root, then mean, then square and then deviation of the variable from its average. It can be written or performed as follows,

$$(v_x^{rms})^2 = < (v_x - < v_x >)^2 > = < v_x^2 > - 2 < v_x < v_x >> + < v_x >^2 = < v_x^2 > - < v_x >^2 . \tag{4.8}$$

We have seen already in the previous chapter, that $< v_x >= 0$, and that $< m v_x^2 > = k_B T$. Hence,

$$v_x^{rms} = \sqrt{\left(\frac{k_B T}{m}\right)}. \tag{4.9}$$

Fig. 4.1 illustrate the distribution of the x-component of particle velocities according to the Maxwell distribution. The root-mean-square velocity v_x^{rms} increases upon increasing temperature, according to Eq. (4.9). Thus, the distribution is wide at a high temperature. In contrast at a low temperatures the distribution is quite narrow.

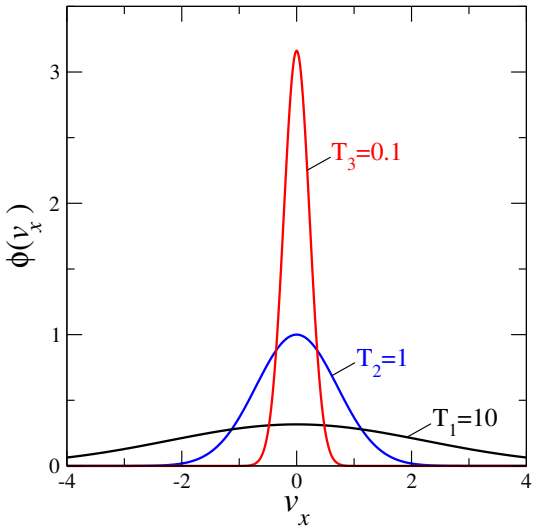


Fig. 4.1: Distribution of the x-component of particle velocities from the Maxwell distribution given by Eq. (4.3) $(m/2\pi k_B = 1)$.

Although the average velocity of a molecule chosen at random belonging to a system taken at random from an ensemble of ideal gases considered at N, V an T conditions is zero as we have obtained above, the average speed of this molecule is non-zero. The

average speed is one of important characteristics of the thermal motion of molecules in the system under consideration. By definition, the speed of a molecule that has velocity $\mathbf{v}$ is the magnitude or the absolute value of velocity, i.e. it is $v = \sqrt{\mathbf{vv}}$.

In order to calculate it, it is convenient to use the spherical coordinates v, θ, φ. Thus, the Cartesian coordinates should be transformed into the spherical coordinates, $\mathbf{v} = \mathbf{i}v_x + \mathbf{j}v_y + \mathbf{k}v_z$ with $v_x = vsin\theta cos\varphi$, $v_y = vsin\theta \sin\varphi$ and $v_z = vcos\theta$. Note that, $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$, and the Jacobian of transition from Cartesian to spherical coordinates to perform next the integration is,

$$dv_x dv_y dv_z = Jdvd\theta d\varphi = v^2 dvsin\theta d\theta d\varphi, \tag{4.10}$$

where $0 \leq v \leq \infty$, $0 \leq \theta \leq \pi$, $0 \leq \varphi \leq 2\pi$. Then, the average speed $< v >$ is,

$$< v >= (m/2\pi k_B T)^{3/2} \int_0^{2\pi} d\varphi \int_0^{\pi} sin\theta d\theta \int_0^{\infty} dv v^3 \exp\{-[mv^2/2k_B T]\} = \sqrt{\frac{8k_B T}{\pi m}}. \tag{4.11}$$

This expression permits to establish relation between the average speed of a molecule or atom in a dilute gas and certain characteristic velocity. It appears that $< v >$ is of the order $10^4 cm/s$, i.e of the same order as the speed of sound (the latter can be interpreted as the transmission of density perturbations in air). Thus, the sound propagation is straightforwardly related to molecular motion.

Additional features of the motion of particles in a dilute gas can be obtained by defining the distribution of speeds. We assume that $\Phi(v)dv$ is the probability that the magnitude or absolute value of the speed of an arbitrary particle of a gas is in the interval $(v, v + dv)$, irrespective of the direction of the vector of velocity, $\mathbf{v}$, of a particle, i.e.,

$$\Phi(v)dv = \int_0^{2\pi} d\varphi \int_0^{\pi} sin\theta d\theta v^2 \phi(\mathbf{v}) dv. \tag{4.12}$$

Thus, the corresponding probability density function of speed is,

$$\Phi(v) = 4\pi v^2 \phi(\mathbf{v}). \tag{4.13}$$

Just recall, that $\phi(\mathbf{v})$ depends on the magnitude of $\mathbf{v}$ only.

In order to find the most probable speed, $\tilde{v}$, it is necessary to solve the equation,

$$\left.\frac{d\Phi(v)}{dv}\right|_{v=\tilde{v}} = 0. \tag{4.14}$$

Taking the derivative, one obtains,

$$\tilde{v} = \left(\frac{2k_B T}{m}\right)^{1/2}. \tag{4.15}$$

One can construct the graph of the function $\Phi(v)$ to see that it has the bell-like shape with the maximum at $\tilde{v}$.

Referring to the previous developments in this chapter, Eqs. (4.7)- (4.9), the rms deviation of velocity can be evaluated taking into account that $v^2 = v_x^2 + v_y^2 + v_z^2$. Then,

$$v_{rms} = \{< (\mathbf{v} - < \mathbf{v} >)^2 >\}^{1/2} =< v^2 >^{1/2} = \sqrt{3k_B T/m}. \tag{4.16}$$

It follows then with the sake of analysis, that,

$$\tilde{v} < < v > < v_{rms}. \tag{4.17}$$

In Fig. 4.2, the distribution of speeds, $\Phi(v)$, in units of $(m/2k_BT)^{1/2}$ as a function of speed v is shown. The most probable speed $\tilde{v}$, the average speed $\langle v \rangle$, and the root-mean-square speed $v_{\rm rms}$ are marked in the figure.

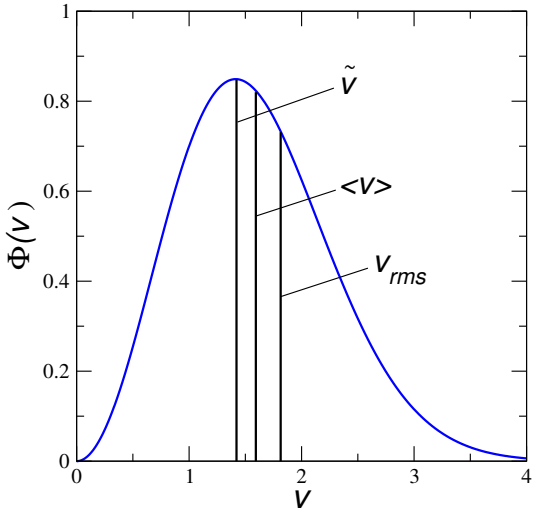


Fig. 4.2: Distribution of speeds, $\Phi(v)$, in units of $(m/2k_BT)^{1/2}$ as a function of speed v.

Having established the sequence of characteristic velocities, let us proceed now to the intent of description of possible collisions between particles. The reason of doing that is in establishing a more comprehensive insight into the properties of dilute gas.

4.2 Mean free path and mean collision frequency

The concepts of the free path and collision frequency represent natural extension of the application of Maxwell distribution of velocities. Evidently, the interactions between particles, atoms or molecules exist in any aggregation state of matter. However, in a dilute gas at equilibrium their effect is negligible, the mean separation of particles is much larger than the interaction range. Still, the inter-particle collisions can happen. In fact, they provide mechanism for reaching the equilibrium state.

In order to characterize collisions, let us introduce the notion of the mean free path, λ. It is the mean distance characterizing motion of a molecule between subsequent collisions. For simplicity, assume that the molecules are rigid spheres of diameter, d. Concerning nature of particles, we would just mention that the collisions observed are perfectly elastic, i.e. without possible dissipation processes. Restrict attention to a one-component gas with the number density $n = N/V$ at temperature T.

Consider a molecule, numbered as 1 with an average speed v_{av} and observe its movement during certain small interval of time δt. Imagine that the molecule is entering a cylinder with the length $v_{av}\delta t$, the radius of the cylinder is taken to be d. (Make a figure on your own for better visualization!). The molecule is moving along the axis of the cylinder. Any molecule with its center inside the cylinder will be struck by the molecule 1 upon passing time δt. Recall, that the diameter of the cylinder is chosen such to permit pass of the molecule 1 and the gas density inside the cylinder is n. The setup assumed, permits to derive an estimate for the number of collision per unit time. Namely, it is the number of molecules that suffered collisions, i.e., $n(v_{av}\delta t)(\pi d^2)$ divided by time δt. Thus, the mean frequency of collisions, ν, is $\pi n v_{av} d^2$. On the other hand, the mean free path, λ then is, $\lambda = v_{av}/\nu$. If we recall that $v_{av} = \sqrt{8k_BT/\pi m}$, then the estimates for ν and λ involve the characteristics of particles, m, d and the characteristics of a gas n, T.

This kind of reasoning is heuristic in the sense that for exploratory purposes we have taken into account only the most important features of the phenomenon under

consideration without referring to any sophisticated mathematical procedure. Specifically, not all the molecules are treated at the same level, the first molecule is moving (the average speed is deduced from Maxwell distribution), whereas all the molecules in the cylinder are assumed motionless. We will see, however, that the truth does not deviate much from such simplified derivation.

Thus, to get a more adequate description, we need to introduce elements of the kinematics of collisions. Pair collisions should be included. What we need is to obtain the probable number of collisions a molecule 1 having velocity in the interval $(\mathbf{v}_1, \mathbf{v}_1 + d\mathbf{v}_1)$ with all molecules that have velocities in the interval $(\mathbf{v}_2, \mathbf{v}_2 + d\mathbf{v}_2)$ during interval of time δt. Therefore, we need to have the following mathematical elements, $\phi(\mathbf{v}_1)d^3v_1$ and $\phi(\mathbf{v}_2)d^3v_2$. However, most important, is the direction of the vector, $\mathbf{v}_{21}$, $\mathbf{v}_{21} = \mathbf{v}_2 - \mathbf{v}_1$, the particles should approach to collide during time interval δt. Consequently, the distance between centers of two particles should be less than $|\mathbf{v}_{21}\mathbf{r}|\delta t$. The unit vector $\mathbf{r}$ is assumed to be directed along line joining the particle 1 and 2.

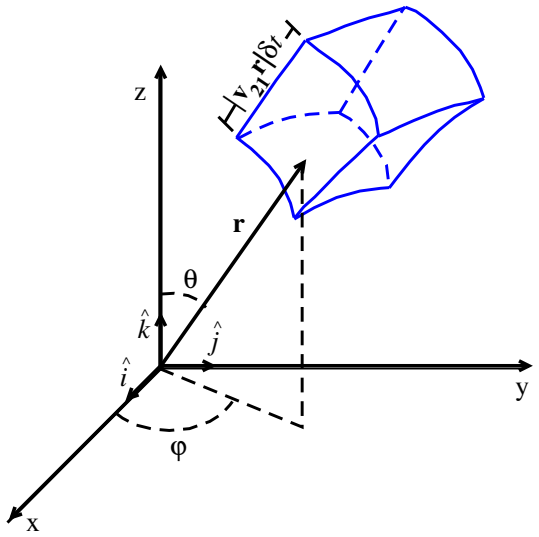


Fig. 4.3: Schematic representation of the volume element within which molecule 2 must lie to collide with molecule 1 within the time interval δt.

Let us choose the coordinate system centered on molecule 1 and the orientation of the z-axis is taken to be along $\mathbf{v}_{21}$ but opposite to the direction of this vector, see Fig. 4.3, (note the direction of the unit vector $\mathbf{k}$). The molecules 1 and 2 will collide if 2 is in the volume element $d\Omega d^2|\mathbf{v}_{21}\mathbf{r}|\delta t$. If θ denotes the angle between the vectors $\mathbf{v}_{21}$ and $\mathbf{r}$, collision will occur if $\pi/2 < \theta < \pi$. Recall that d is the diameter of the molecules and equals to the distance between centers of molecules upon touching each other, $d\Omega$ denotes an element of the solid angle. So, what we need to calculate is the expression,

$$\phi(\mathbf{v}_1)d^3v_1(nd^2\,|\mathbf{v}_{21}\mathbf{r}|\,\delta t d\Omega)\phi(\mathbf{v}_2)d^3v_2. \tag{4.18}$$

The average collision frequency, ν, is the number of collisions per time interval. Moreover, we need to integrate the expression above over $\mathbf{v}_1, \mathbf{v}_2$ and over the solid angle. Solely, the additional constraint, $\mathbf{v}_{21}\mathbf{r} < 0$, or equivalently $\pi/2 < \theta < \pi$, should be taken into account to provide collision. Then, necessary expression to deal with is,

$$\nu = \begin{cases} nd^2 \int d^3v_1 \int d^3v_2 \int d\Omega\phi(\mathbf{v}_1)(|\mathbf{v}_{21}\mathbf{r}|)\phi(\mathbf{v}_2) = \\ \\ -2\pi nd^2 \int\limits_{\pi/2}^{\pi} cos\theta sin\theta d\theta \int d^3v_1 \int d^3v_2 |\mathbf{v}_{21}|\phi(\mathbf{v}_1)\phi(\mathbf{v}_2) = \\ \\ \pi nd^2 \int d^3v_1 \int d^3v_2 |\mathbf{v}_{21}|\phi(\mathbf{v}_1)\phi(\mathbf{v}_2). \end{cases} \tag{4.19}$$

The final step of calculations requires transformation of velocity coordinates. Namely, it is necessary to make transformation from $\mathbf{v}_1$ and $\mathbf{v}_2$ to the sum and difference,

$$\mathbf{G} = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2), \quad \mathbf{g} = (\mathbf{v}_2 - \mathbf{v}_1). \tag{4.20}$$

Then,

$$v_1^2 + v_2^2 = 2G^2 + g^2/2. \tag{4.21}$$

Taking into account that, $d^3v_1 d^3v_2 = d^3G d^3g$, then substituting the Maxwell distributions explicitly, and using spherical coordinates, the last line of Eq. (4.19) becomes:

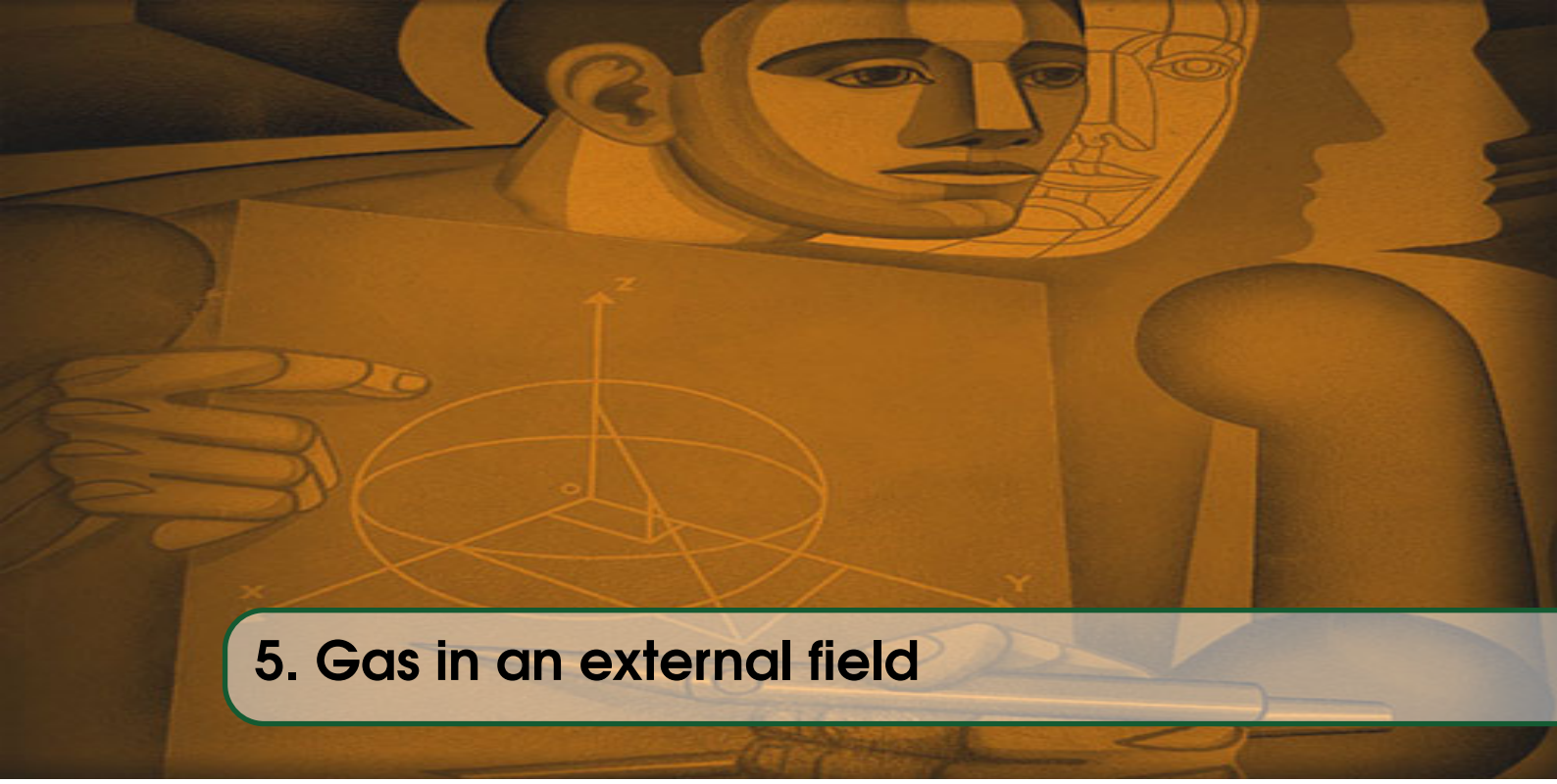
$$\nu = \pi n d^2 (m/2\pi k_B T)^3 (4\pi)^2 \int_0^\infty dG G^2 \exp\left[-mG^2/kT\right] \int_0^\infty dg g^3 \exp\left[-mg^2/4kT\right]. \tag{4.22}$$

The integrals of this sort has been used in the previous chapters. Then, the final result is,

$$\nu = \sqrt{2} n \left(\frac{8k_B T}{\pi m}\right)^{1/2} \pi d^2, \qquad \lambda = \frac{1}{(\sqrt{2})\pi n d^2}. \tag{4.23}$$

These two expressions show what kind of corrections enter the results coming out from heuristic arguments.

Exercise 4.1 Perform integrations explicitly in Eq. (4.22) and obtain the result. ■



5. Gas in an external field

5.1 Boltzmann factor and barometric formula

Our developments in the previous chapters have been restricted to homogeneous dilute gases, any kind of external influence or external field has not been present. Thus, the density of particles is constant throughout the system. However, in the presence of the external field, the situation becomes different. Evidently, the air density at Acapulco is different from its density at the top of Mount Everest. It is the result of action of the gravitational field.

Let us consider, this kind of problem from the fundamental point of view. The system in question is a dilute isothermal gas in the presence of the gravitational field. The force acting on particles is assumed conservative, i.e. it is determined by the scalar potential, $u(\mathbf{r})$, and is $\mathbf{F} = -\nabla u(\mathbf{r})$. If the force acting on the particle is due to gravitational field, it is $\mathbf{F} = -mg\mathbf{k}$, where $\mathbf{k}$ is the unit vector normal to the earth's surface pointing vertically upwards. The potential generating this force is $u(\mathbf{r}) = mg\mathbf{k}\mathbf{r} = mgz$, where z is the coordinate of the particle in this frame.

We would like to find the relation between the local hydrostatic pressure, $P(r)$, of a fluid at equilibrium, and the external force $\mathbf{F}(\mathbf{r})$. To do that, consider a small volume element of a fluid at arbitrary point of space (x,y,z), with the volume $\Delta x\Delta y\Delta z$. The pressure of the fluid surrounding the element leads to a surface force acting on each face of the volume element. This force is normal to the face and is directed into the element. In the Cartesian coordinates in laboratory frame with z axis pointed upwards, the net force acting from outside into the element is,

$$-\mathbf{k}\int_{x}^{x+\Delta x} dx' \int_{y}^{y+\Delta y} dy' [P(x',y',z+\Delta z) - P(x',y',z)], \tag{5.1}$$

along z-axis, and,

$$-\mathbf{j}\int_{x}^{x+\Delta x} dx' \int_{z}^{z+\Delta z} dz' [P(x',y+\Delta y,z') - P(x',y,z')], \tag{5.2}$$

$$-\mathbf{i}\int_{y}^{y+\Delta y} dy^{'}\int_{z}^{z+\Delta z} dz^{'}[P(x+\Delta x,y^{'},z^{'})-P(x,y^{'},z^{'})], \tag{5.3}$$

along y and x axes, respectively. Fig. 5.1 shows a schematic representation of body and surface forces acting on a small cubic volume element of a fluid at equilibrium.

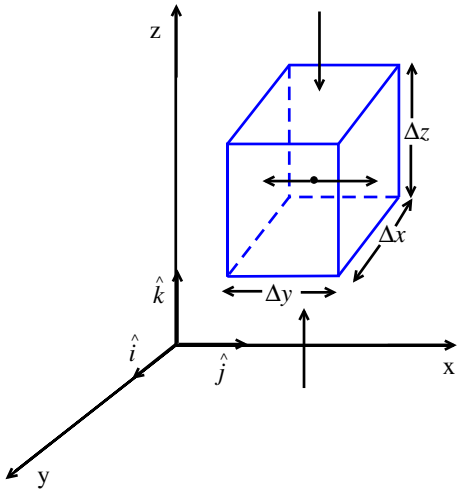


Fig. 5.1: Schematic representation of body and surface forces on a small cubic box in an equilibrium fluid.

The balance between the surrounding and the body also involves the opposite force. The body force, in contrast to the surface force, expresses action of the body from the inside of the volume element to outside. The fluid in the volume has the number density n. Thus, the body force can be formally written as follows,

$$\int_{x}^{x+\Delta x} dx^{'}\int_{y}^{y+\Delta y} dy^{'}\int_{z}^{z+\Delta z} dz^{'}[\mathbf{i}nF_x(x^{'},y^{'},z^{'})+\mathbf{j}nF_y(x^{'},y^{'},z^{'})+\mathbf{k}nF_z(x^{'},y^{'},z^{'})]. \tag{5.4}$$

Because of hydrostatic equilibrium, the sum of two forces should cancel or in other words the force balance should fulfill. Hence, we can write down three equations expressing this fact. In the case of x-component, for example, we have,

$$\begin{aligned}&-\frac{1}{\Delta y\Delta z}\int_{y}^{y+\Delta y} dy^{'}\int_{z}^{z+\Delta z} dz^{'}\frac{[P(x+\Delta x,y^{'},z^{'})-P(x,y^{'},z^{'})]}{\Delta x}\\&+\frac{1}{\Delta y\Delta z\Delta x}\int_{x}^{x+\Delta x} dx^{'}\int_{y}^{y+\Delta y} dy^{'}\int_{z}^{z+\Delta z} dz^{'}nF_x(x^{'},y^{'},z^{'})=0,\end{aligned} \tag{5.5}$$

where the force balance equation has been divided by the volume of the element. Let us perform the limit $\lim_{\Delta x\to 0}$, to deal with infinitesimal volume. Then, the integrand in the first, upper line of equation yields $\partial P(x,y^{'},z^{'})/\partial x$. Concerning the second, lower line in Eq. (5.5), it is,

$$\lim_{\Delta x\to 0}\int_{x}^{x+\Delta x} dx^{'}nF_x(x^{'},y^{'},z^{'})=nF_x(x,y^{'},z^{'}), \tag{5.6}$$

in accordance to the mean value theorem. In consequence, the equation above becomes,

$$\frac{1}{\Delta y \Delta z} \int_{y}^{y+\Delta y} dy^{'} \int_{z}^{z+\Delta z} dz^{'} \left[-\frac{\partial P(x,y^{'},z^{'})}{\partial x} + nF_x(x,y^{'},z^{'}) \right] = 0. \tag{5.7}$$

Next, we would like to perform successively two more limits, $\lim_{\Delta y \to 0}$ and $\lim_{\Delta z \to 0}$, using the mean value theorem, and obtain,

$$-\frac{\partial P(x,y,z)}{\partial x} + nF_x(x,y,z) = 0. \tag{5.8}$$

Similar procedure should be performed w.r.t. the y and z components of the force balance equation. With these three results if combined and multiplied by unit vectors of the Cartesian coordinate system , the condition of hydrostatic equilibrium leads to the following result,

$$-\nabla P + n\mathbf{F} = 0. \tag{5.9}$$

It is worth noting that we have not resorted to the dilute gas concept explicitly. Exploit this notion now by using the equation of state of ideal gas, $P = nk_BT$ and recall the definition of the force, $\mathbf{F} = -\nabla u$. Then,

$$-\nabla P - \frac{P}{k_BT}\nabla u = 0; \quad \nabla(lnP) = -\frac{1}{k_BT}\nabla u. \tag{5.10}$$

Interestingly, we observe that this equation expresses spatial dependence of the pressure of a dilute gas in terms of the potential energy of the system due to the external field solely (the potential energy due to inter-particle interactions is zero within the ideal gas model). Moreover, we have not resorted to the notion of gravitational field in the procedure. Thus, the result in our disposal is general, it applies to any kind of conservative forces related to the scalar potential of an external field.

The solution of the first-order differential Eq. (5.10) has the form,

$$P(\mathbf{r}) = P(\mathbf{r}_0) \exp\left\{ -\frac{[u(\mathbf{r}) - u(\mathbf{r}_0)]}{k_BT} \right\}, \tag{5.11}$$

where $P(\mathbf{r}_0)$ and $u(\mathbf{r}_0)$ represent the boundary condition describing the pressure and the potential energy due to the external field at some reference point $\mathbf{r}_0$, respectively.

In the particular case, when the external filed is the gravitational field, the result is,

$$P(z) = P(z_0) \exp\left\{ -\frac{mg}{k_BT}(z - z_0) \right\}, \tag{5.12}$$

known as the barometric formula describing change of pressure with relative height $z - z_0$ for isothermal system.

There exists another line of analysis of the final result for pressure for the system under the influence of the external field.

5.2 Analysis and consequences

One can exclude pressure, P, from Eq. (5.10) in favor of fluid density for ideal gas ($n = P/k_BT$), to arrive at the Boltzmann equation,

$$n(\mathbf{r}) = n(\mathbf{r}_0) \exp\left\{ -\frac{[u(\mathbf{r}) - u(\mathbf{r}_0)]}{k_BT} \right\}. \tag{5.13}$$

The ratio, $n(\mathbf{r})/n(\mathbf{r}_0)$, describes the relative number density of particles at a point w.r.t. the point of reference. It is determined by the Boltzmann factor, $\exp\{-[u(\mathbf{r}) - u(\mathbf{r}_0)]/k_BT\}$.

Commonly, the Boltzmann factor is the exponential, containing either the interaction potential between particles or the potential describing the external fields as we have at present.

Interestingly, the equation permits seducing statistical interpretation. If we denote by d^3r the volume element at $\mathbf{r}$, and interpret $n(\mathbf{r})$ as the distribution of the number density of particles as in fact it is, then $n(\mathbf{r})d^3r$ is the probable number of particles in this volume interval, i.e. in the interval $(\mathbf{r}, \mathbf{r}+d^3r)$. The dilute gas in question has been considered at under N, V and T conditions. Hence,

$$N = \int\limits_V n(\mathbf{r})d^3r = n(\mathbf{r}_0)\exp\{u(\mathbf{r}_0)/k_BT\}\int\limits_V \exp\{-u(\mathbf{r})/k_BT\}d^3r. \tag{5.14}$$

Therefore, we are able now to eliminate unknown constant $n(\mathbf{r}_0)\exp\{u(\mathbf{r}_0)/k_BT\}$,

$$n(\mathbf{r}_0)\exp\{u(\mathbf{r}_0)/k_BT\} = N/\int\limits_V \exp\{-u(\mathbf{r})/k_BT\}d^3r. \tag{5.15}$$

The number density distribution given by Eq. (5.13) can be written thus in a more elegant form,

$$n(\mathbf{r}) = N\frac{\exp\{-u(\mathbf{r})/k_BT\}}{\int\limits_V \exp\{-u(\mathbf{r})/k_BT\}d^3r}. \tag{5.16}$$

Let, introduce now the function $p(\mathbf{r})$ by the definition,

$$p(\mathbf{r}) = \frac{n(\mathbf{r})}{N} = \frac{n(\mathbf{r})}{\int\limits_V n(\mathbf{r})d^3r}. \tag{5.17}$$

According to the comments above, concerning the meaning of $n(\mathbf{r})d^3r$, we observe that the function $p(\mathbf{r})$ has sense of the probability density or the distribution function in the coordinate space, as it is correctly normalized $\int\limits_V d^3r p(\mathbf{r}) = 1$. However, this function refers to a particle in the N-particle system.

As we have seen in the first chapter, a microstate of a system of N particles requires to specify their coordinates and velocities (or moments). If our interest is in a single particle, we can combine present knowledge and our previous development in terms of Maxwell distribution. So, if we would like to find the probability for a particle to be in a volume d^3r at a coordinate in the interval $(\mathbf{r}, \mathbf{r}+d^3r)$ and with velocity in the interval $(\mathbf{v}, \mathbf{v}+d^3v)$ it would be as follows,

$$p(\mathbf{r})d^3r\phi(\mathbf{v})d^3v = C\exp\{-[mv^2/2 + u(\mathbf{r})]\}d^3vd^3r, \tag{5.18}$$

where C is a constant.

From the point of view of ensembles of the systems with N particles, it is seducing and suggestive to construct the distribution function describing configurations of N particles such that each of them has certain velocities. Namely, we assume that the desired function will have the form,

$$P_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N; \mathbf{v}_1, \mathbf{v}_2, .., \mathbf{v}_N) = C^N\exp\{-\sum_{i=1}^{N} mv_i^2/2 + U_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N)\}. \tag{5.19}$$

Within the model of ideal gas, the particles are independent, the external field acts on each of them at their independent positions, thus, the potential energy of the system consists of independent contributions,

$$U_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N) = \sum_{i=1}^{N} u(\mathbf{r}_i). \tag{5.20}$$

However, if the particles interact between themselves (as in non-ideal systems), this function is more complex,

$$U_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N) = \sum_{i=1}^{N} u(\mathbf{r}_i) + \sum_{i,j(i<j)}^{N} u_2(\mathbf{r}_i, \mathbf{r}_j), \tag{5.21}$$

where $u_2(\mathbf{r}_i, \mathbf{r}_j)$ is the pair potential of inter-particle interaction. This kind of systems using realistic models will be the subject of chapters in the final part of our course of statistical thermodynamics. Moreover, the distribution function of the form given by Eq. (5.19) will be derived by using the theory of statistical ensembles.

6. Intermolecular pair potential

6.1 Intermolecular interactions

As we have mentioned in the previous chapter, the potential energy of the system of interacting particles as a function of coordinates has rather complex form. In the case of the system of one species it can be written as,

$$U_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N) = \sum_{i=1}^{N} u(\mathbf{r}_i) + \sum_{i,j(i<j)}^{N} u_2(\mathbf{r}_i, \mathbf{r}_j), \tag{6.1}$$

where $u_2(\mathbf{r}_i, \mathbf{r}_j)$ is the pair potential of inter-particle interaction. The first contribution comes in play, if the system is under the influence of an external field. On the other hand, the expression given by Eq. (6.1) is incomplete, because formally, there exists an interaction between three particles, $u_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k)$, and of the higher order. This kind of contributions, coming from many-body interactions, into thermodynamic properties are eventually quite small and are neglected in the overwhelming majority of problems in the theory of fluids. Thus, in the entire text of these chapters, we restrict to the form given by Eq. (6.1). Still, to be more comprehensive, it is necessary to mention that the contribution from pair interaction into the potential energy of many-component systems is,

$$\sum_{a,b}^{M} \sum_{i,j(i<j)}^{N} u_{ab}(\mathbf{r}_i, \mathbf{r}_j), \tag{6.2}$$

where M denotes the number of kinds (sorts) of species in the system. If one restricts to homogeneous and isotropic fluids, the inter particle potential depends on the distance between particles only,

$$u_2(\mathbf{r}_i, \mathbf{r}_j) = u_2(|\mathbf{r}_i - \mathbf{r}_j)|). \tag{6.3}$$

The theory of inter-molecular forces represents a quite complicated application of the principles of quantum mechanics requiring consideration of an assembly of nuclei and electrons of atoms or molecules. This issue is out of scope of the present text. Rather we will resort to various models of inter-molecular interactions to capture principal or most important features of the behavior of specific fluid systems of many particles. In other words, we will use as realistic models as possible for systems in question.

One of the important features of model potentials used in the studies of liquids and solutions is that the pair interaction between particles is repulsive at small inter-particle distances. This repulsion arises from a rather sophisticated quantum mechanical considerations that are out of scope of our lecture notes.

The range of repulsion is of the order of average distance between neighboring particles. The attractive interactions act at larger distances, they change with distance moderately and contribute less to the liquid structure. In certain sense, we can say that short-range interaction contribute mostly to the entropic effects whereas the attractive interactions principally make effect to energetic properties. Evidently, the system of interacting particles is unable to convert into liquid from its vapor without inter-particle attraction.

Various theoretical developments in the liquid state theory are based on the perturbation treatment within which the reference system or "model" contains solely the repulsive interaction. Such simplest and useful model is the fluid of hard spheres in which the pair interaction is,

$$u_2(r) = \begin{cases} \infty, & r < d \\ 0, & r > d \end{cases}, \tag{6.4}$$

where d is the diameter of spherical particles. Within this model, upon compression a dilute system will convert into a dense system monotonously, i.e. there is no distinction between say vapor and liquid. Consequently, one would expect to arrive at either glass-like or solid phase at a high pressures. The hard sphere fluid model permits development of several analytic and semi-analytic theories and the application of computer simulations methodology (like Monte Carlo or discontinuous molecular dynamics) as well. The model of hard spheres is characterized solely by a single parameter given as a reduced dimensionless density, $\rho^* = Nd^3/V$, or as a packing fraction, $\eta = V_{occ}/V = \pi N d^3/6V$ describing ratio of the volume occupied by particles w.r.t. the total volume of the system. Maximal packing, η_{max}, depends on the system dimensionality. Evidently, $\eta_{max} = 1$ for 1D fluid. For 2D fluid (fluid of hard spheres confined to a plane) $\eta_{max} = \pi\sqrt{3}/6 \approx 0.9069$. Finally, in 3D $\eta_{max} = \pi\sqrt{2}/6 \approx 0.7405$.

Exercise 6.1 Analyze the structure and maximum number of neighbors, demonstrating that the maximum packing fraction η_{max} for a 2D fluid and a 3D fluid is approximately $\eta_{max} \approx 0.9069$ and $\eta_{max} \approx 0.7405$, respectively. Additionally, examine the three most common crystal structures: i) body-centered cubic (bcc), ii) face-centered cubic (fcc), and iii) hexagonal close-packed (hcp). Finally, discuss the significance of the freezing density for the hard-sphere model at $\eta = 0.494$ and the glass transition density, estimated at $\eta \approx 0.6$, and provide an explanation for this difference. ■

Possible, and frequently used extension of this kind of potential is the so-called square-well fluid model with the potential in the form,

$$u_2(r) = \begin{cases} \infty, & r < d \\ -\varepsilon, & d < r < \lambda d \\ 0, & r > \lambda d \end{cases}, \tag{6.5}$$

where λ and ε are the width and depth of attractive interaction, respectively. This model potential is appealing because it permits analytic procedure to be applied. Usually, the width of the attractive interaction should be chosen neither too wide nor too narrow, e.g. between 1.5 and 2.0. On the other hand, relation between the potential energy of interaction with respect to thermal energy is described by the parameter ε/k_BT. Thus, one has two parameters for such model, the packing fraction, η, and reduced dimensionless temperature, $T^* = k_BT/\varepsilon$. The model permits phase transition from vapor into liquid and next to a solid phase, dependent on the values of the parameters.

However, more realistic representation of the interaction between a pair of uncharged

particles is the so-called (12-6) Lennard-Jones (LJ) potential of the form,

$$u_2(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{6.6}$$

where σ and ε are the diameter of the molecules and the depth of their attraction, respectively. The potential vanishes as $r = \sigma$ and attains minimum at $r = 2^{1/6}\sigma$. This potential is adequate to model noble gases and methane as well. Examples of various inter-particle potentials, including the widely-used Lennard-Jones pair potential are shown in Fig. 6.1.

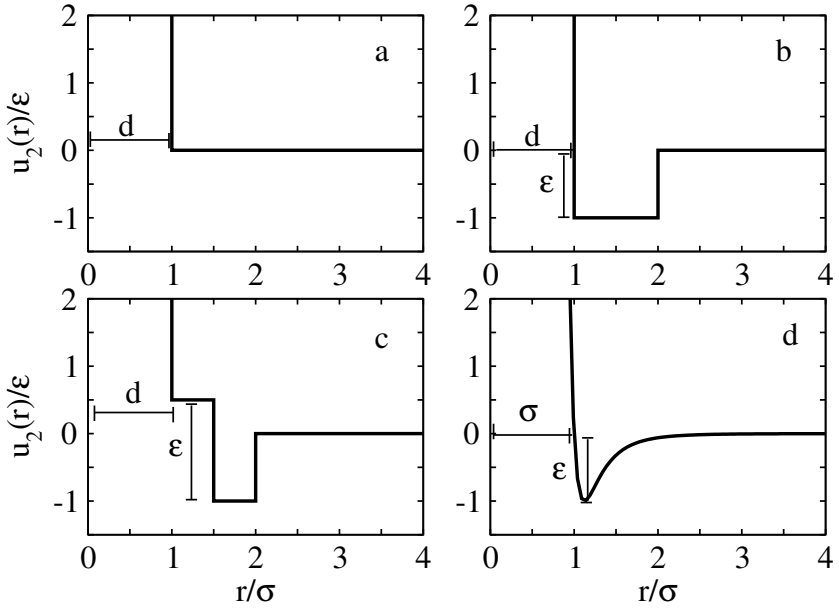


Fig. 6.1: Examples of short-range inter-particle potentials. Illustrated pair potentials include: (a) hard-sphere (see Eq. (6.4)), (b) square-well (see Eq. (6.5)), (c) square-well – square-shoulder, and (d) Lennard-Jones (see Eq. (6.6)).

Several key factors should be considered when describing inter-particle potentials:

i) The particles we deal with not necessarily should be spherical. Namely, if one aims in the description of liquid crystals, both of the microscopic structure and thermodynamics, the Gay-Berne model can be used. It is an anisotropic form of the Lennard-Jones 12-6 potential. Detailed description of the potential and the results for the phase diagram can be found in the Ref. [11].

ii) We would like to recommend additional reading concerning the modelling of water by using non-electrostatic potentials, see e.g. Ref. [12, 13].

Let us proceed now to long-ranged potentials. The fundamental interaction between charged particles is given by Coulomb potential. Let us write down it for the case of two point particles with charges ez_i and ez_j separated by the distance r,

$$u_{ij}(r) = e^2 z_i z_j \frac{1}{r}, \tag{6.7}$$

where e is the elementary charge. It is the central type potential, i.e. the force is acting along the line between centers of the particles (the force is proportional to $1/r^2$). Moreover, it is slowly decaying potential, in comparison to other potentials of the type, $1/r^n$ with $n > 1$. The potential is attractive if two charges possess opposite charge and repulsive otherwise. One comment concerns electrostatic units. If one works in SI units, the proportionality constant should be $1/4\pi\varepsilon_0$, where ε_0 is the vacuum electric permittivity (the all books uses the CGS units). One should have in mind that ions possess certain diameters, therefore in the formulation of the theory of electrolyte solutions or ionic melts, one must combine

e.g. the hard sphere potential describing short-range interactions for $r < d$ and long-range Coulomb interactions. At the level of primitive modelling of electrolyte solutions only ionic solutes are taken into account explicitly. The solvent is considered as a uniform structureless medium characterized by a single parameter ε, it is the dielectric constant. Then, it is assumed that ions interact via Coulomb potential with account of implicit solvent, $u_{ij}(r) = e^2 z_i z_j / \varepsilon r$.

Our final comment at this point is the following. The system of charges should be electroneutral in total to provide an equilibrium state, i.e. should satisfy the condition,

$$\sum_{a=1}^{M} e z_a N_a = 0, \tag{6.8}$$

where M is the number of types of ions, z_a is the valency of ions of species a and N_a is the number of ions of species a.

Exercise 6.2 How to make Fourier-transform of the Coulomb potential?. ■

Proceed to one more interesting and important concept. A neutral molecule can be viewed as a charged nucleus surrounded by electron cloud. If a charge interacts with the electron cloud, it will deform it, better say polarize it. The interaction potential between a charge with the magnitude ez_i and the polarized molecules is,

$$u(r) = -e^2 z_i^2 \frac{\alpha}{r^4}. \tag{6.9}$$

The parameter α is called the polarizability of the molecule. Actually, it is of the order of the molecular volume. The potential is of central type (if the molecule is spherical), moreover it is short-ranged ($\propto r^{-4}$). Also, it is worth mentioning that this potential does not depend on the sign of charge of the particle that polarizes a molecule, i.e. polarizability is the characteristics of the molecule under consideration. Finally, if the molecule is of non-spherical shape, then the polarization potential is angle-dependent, it will depend on the angle between two vectors, one directed along inter-particle distance between centers, and another along the major axis of the molecule

Now consider one more physical situation. If in a neutral molecule the nuclear and electronic charge distribution is such that there is a separation of the centers of positive and negative charge, then the molecule possesses a permanent dipole moment, $\boldsymbol{\mu}$, it has dimension of charge times distance between charge distributions. It is measured in Debye units. In general,

$$\boldsymbol{\mu} = \sum_i q_i \mathbf{r}_i, \tag{6.10}$$

where q_i describe charge distribution and $\mathbf{r}_i$ are defined accordingly in the center of mass (COM) of molecule frame. In Fig.6.2, a schematic representation of water and acetone molecules illustrating the direction of their dipole moments is given.

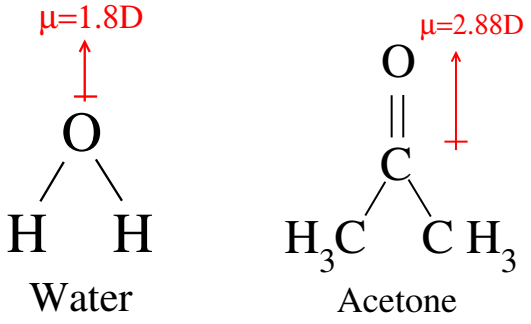


Fig. 6.2: Schematic representation illustrating the direction of the dipole moment of water and acetone molecules.

For purposes of statistical thermodynamics of electrolyte solutions or melts, or ion-molecular plasma, one needs the potential of interaction between ions (or just say charges) and a molecule having permanent dipole moment. It has the following form,

$$u_{i-d}(\mathbf{r},\boldsymbol{\mu}) = -ez_i\frac{\boldsymbol{\mu}\mathbf{r}}{r^3}. \tag{6.11}$$

Evidently, this model potential should be applied at not very small inter-particle distance, at least outside of the cores describing each particle. The potential decays as r^{-2} ($\mathbf{r}/r$ defines the unit vector direction) so this potential is long-ranged. It depends on the distance between the charge and the COM of a molecule, and on the angle between the dipole moment vector $\boldsymbol{\mu}$ and vector $\mathbf{r}$. In several applications the true dipole moment is substituted by a point dipole (of zero length) but the orientation is preserved.

Fluids of molecules that have permanent dipole moment are referred to as polar fluids. Water, alcohols, dimethylsulfoxide are illustrative examples. The dipole-dipole interaction is defined as follows,

$$u_{d-d}(\mathbf{r},\boldsymbol{\mu}_i,\boldsymbol{\mu}_j) = -\frac{\boldsymbol{\mu}_i\boldsymbol{\mu}_j r^2 - 3(\boldsymbol{\mu}_i\mathbf{r})(\boldsymbol{\mu}_j\mathbf{r})}{r^5}. \tag{6.12}$$

The dipole-dipole potential as a function of inter-particle distance decays as r^{-3}. This type of decay actually is used to distinguish formally between short-range and long-range potentials. However, the dipole-dipole interaction depends on the orientations of two dipoles, $\boldsymbol{\mu}_i,\boldsymbol{\mu}_j$. The corresponding angles as well as the orientation of $\mathbf{r}$, are given in laboratory frame.

6.2 Conversion of the pair interaction potential into external field

We would like to derive the expression for the gas-solid potential energy in the form of external field starting from pair interaction potentials. The following derivation is presented according to original development by W. A. Steele [14, 15]. Consider the following problem.

Put a gas in contact with a solid. Suppose a pair-wise interaction potential between a gas atom and any atom of the solid, $u_{gs}(\rho_i)$, is given in the form of the Lennard-Jones (12-6) potential,

$$u_{gs}(\rho_i) = 4\varepsilon_{gs}\left[\left(\frac{\sigma_{gs}}{\rho_i}\right)^{12} - \left(\frac{\sigma_{gs}}{\rho_i}\right)^{6}\right], \tag{6.13}$$

where ρ_i denotes the distance between the gas atom relative to the i-th atom in the solid, schematically shown in Fig. (6.3).

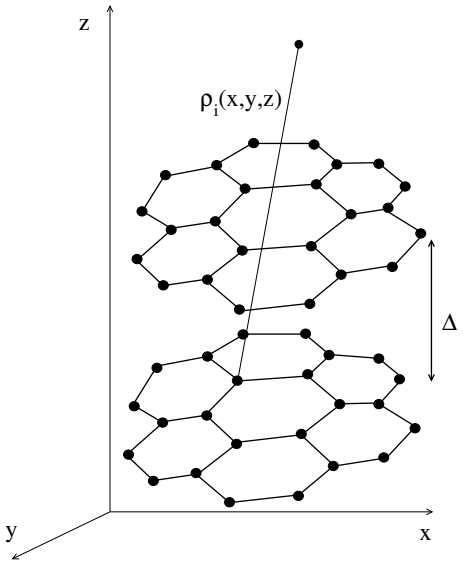


Fig. 6.3: Schematic representation of gas atom over a fragment of graphite solid in the laboratory frame. Δ denotes the spacing between carbon layers.

The parameters ε_{gs} and σ_{gs} describe the depth of the attractive energy well and the size parameter (cross diameter) of the interaction between gas and solid atoms, respectively. It is assumed that the total gas-solid energy, $U_s(\mathbf{r})$, is equal to the sum of pair-wise energies,

$$U_s(\mathbf{r}) = \sum_i u_{gs}(\rho_i). \tag{6.14}$$

Here, $\mathbf{r}$ denotes the position of the gas atom relative to some convenient point in the solid body and the sum in the Eq. (6.14) is performed over all atoms of the solid. So, what we wish to do, is to obtain an expression for energy of interaction of a gas atom with the entire solid.

The principal idea of calculations and the principal approximation is in replacing the discrete set of interaction sites in the solid by a continuous distribution. In other words, we would like to replace summation in Eq. (6.14) by integration. Let us work in Cartesian coordinates. Denote the perpendicular distance between the gas atom and the solid surface as z. Put the solid surface in the xy plane at $z = 0$. Replace the discrete variable ρ_i by the continuous variable ρ. It is defined as,

$$\rho = \sqrt{(z+Z)^2 + X^2 + Y^2}, \tag{6.15}$$

where we denote the coordinates of an arbitrary volume element in the solid as $z + Z$ and X, Y. Then, our aim is to obtain,

$$U_s(z) = \int\limits_{solid} u_{gs}(\rho) n dV, \tag{6.16}$$

that expresses the energy of a gas atom w.r.t. the entire solid as a function of the distance of this gas atom along normal to the solid surface located at $z = 0$. Evidently, it is convenient to use cylindrical coordinates as we have circular symmetry in the surface plane. Note, that the number of interacting sites in the unit volume of the solid (i.e. the solid density) is denoted by n just above. At this stage let us make use of Eq. (6.13) and Eq. (6.15) and

substitute them into Eq. (6.16). Moreover, take into account that $dV = dXdYdZ$, for X and Y use polar coordinates, i.e. $X^2 + Y^2 = s^2$, $dXdY = sdsd\varphi = \pi d(s^2)$. Introduce convenient notation $S = s^2$. Then, Eq. (6.16) takes form,

$$U_s(z) = 4\varepsilon_{gs}\pi n \int_0^\infty dZ \int_0^\infty dS \left\{ \frac{\sigma_{gs}^{12}}{[(z+Z)^2+S]^6} - \frac{\sigma_{gs}^{6}}{[(z+Z)^2+S]^3} \right\} \tag{6.17}$$

Next, simple integration over S variable can be performed to yield,

$$U_s(z) = 2\pi n\varepsilon_{gs} \int_0^\infty dZ \left\{ \frac{2\sigma_{gs}^{12}}{5(z+Z)^{10}} - \frac{\sigma_{gs}^{6}}{(z+Z)^4} \right\}. \tag{6.18}$$

Final integration leads to the important result,

$$U_s(z) = \frac{2}{3}\pi n\varepsilon_{gs}\sigma_{gs}^3 \left\{ \frac{2}{15}\left(\frac{\sigma_{gs}}{z}\right)^9 - \left(\frac{\sigma_{gs}}{z}\right)^3 \right\}. \tag{6.19}$$

It is the interaction potential of a single gas atom with the entire solid as a function of z. In other words, the solid exerts the external field acting on the gas atom that depends on the distance of the gas atom from the solid along normal to the surface plane. The parameters entering this equation have transparent physical meaning. A realistic approximation for $U_s(z)$ in the case of graphite solid is the frequently used 10-4-3 gas-solid interaction potential [14, 15]. It is given as,

$$U_s(z) = 2\pi n\sigma_{gs}^3\varepsilon_{gs} \left\{ \frac{2}{5}\left(\frac{\sigma_{gs}}{z}\right)^{10} - \left(\frac{\sigma_{gs}}{z}\right)^4 - \frac{\sigma_{gs}^4}{3\Delta(z+0.61\Delta)^3} \right\}, \tag{6.20}$$

where Δ represents the spacing between carbon layers, cf. Fig. (6.3).

In order to obtain the cross-diameter and cross interaction energy parameter one needs to apply the Lorentz-Berthelot (LB) combination rules, for example. They refer to the equations that determines the cross-diameter and cross energy of interaction of particles belonging to different species,

$$\sigma_{gs} = \frac{1}{2}(\sigma_{gg} + \sigma_{ss}), \tag{6.21}$$

where σ_{ss} and σ_{gg} are the diameters of the solid-solid and gas-gas interaction, respectively. On the other hand, the cross energy with LB rules, is determined by the geometric mean,

$$\varepsilon_{gs} = \sqrt{\varepsilon_{gg}\varepsilon_{ss}}, \tag{6.22}$$

where ε_{ss} and ε_{gg} are the energy parameters of the solid-solid and gas-gas interaction, respectively. Other types of rules also have been proposed previously and used in computer simulations of liquid mixtures as well [16]. Fig. 6.4 illustrates the gas-solid energy of interaction on the distance from the solid surface. The strength of fluid-solid attraction is the crucial parameter in adsorption studies. It determines if a monolayer or a film of several fluid layers is formed upon adsorption on a solid [17].

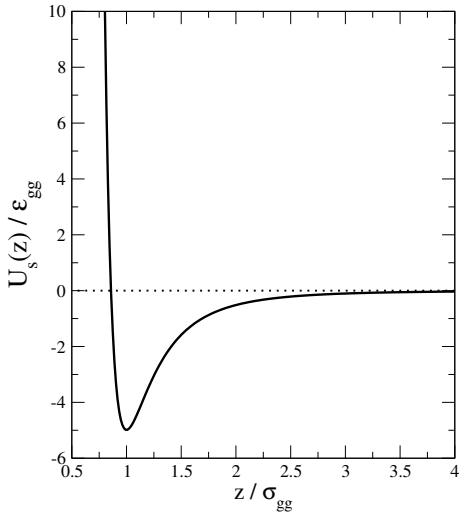


Fig. 6.4: The 10-4-3 gas-solid energy potential.

Exercise 6.3 In Fig. (6.4), we show the gas-solid energy interaction, $U_s(z)$, for a gas in contact with graphite. The parameters for a graphite surface [14, 15] are the following:

- Solid density, ρ_s =114 nm^{-3}.
- Carbon layer spacing between two-parallel layers of graphite, $\Delta = 0.335$ nm.
- Solid-solid parameters for a graphite surface: σ_{ss} =0.340 nm, and ϵ_{ss}/k_B =28 K.
- The fluid-fluid parameters for water are: σ_{gg} =0.30342 nm, and ϵ_{gg}/k_B =250 K.

Using these parameters write the computer program and reproduce the graph presented in Fig. (6.4).

■

7. Quantum mechanics

7.1 Degeneracy temperature

Preliminary comments of this chapter concern with the following question. Previous developments have been concerned with systems consisting of particles that obey the laws of motion of classical mechanics. Classical approximation is valid as a limiting case of a more precise and more universal quantum theory. So, under which conditions a system can be described by classical mechanics?

We begin with a common equation related to the uncertainty principle in quantum mechanics, $\Delta q \Delta p \propto h$. Here Δq and Δp denote uncertainty of the coordinate and of momenta. Higher precision in the evaluation of one variable leads to a lower precision in describing another variable. Let us refer to a particle i of a system of many particles with density $n = N/V$. Assume that its coordinate is q_i and its momentum is p_i. In classical mechanics there is no law prohibiting simultaneous measurement of the coordinate of a particle (body) and its velocity (or momentum). The characteristic length of the fluid system in question is the average distance between particles $< r >$. Thus, to have classical mechanics valid (or to have negligible effect from uncertainties), one can formulate the following conditions,

$$\Delta q_i << \ < r >, \quad \Delta p_i << p_i. \tag{7.1}$$

The momentum of a particle can take values between 0 and ∞. However, if the average momentum is not too small, i.e. the system is not too cold, then there are not many particles with small momenta, thus we can modify the second condition by the following relation,

$$\Delta p_i << \ < p > . \tag{7.2}$$

In consequence, these two conditions, if applied, lead to the relation,

$$\Delta q \Delta p << \ < r >< p > . \tag{7.3}$$

The minimal value for the l.h.s. is the Planck constant h according to the Heisenberg uncertainty principle. It is a fundamental principle of quantum mechanics that establishes a limit of precision for simultaneous determination of coordinate and momentum of a particle. It should be valid for the average values of a coordinate and momentum,

$$h << \ < r >< p > . \tag{7.4}$$

In quantum mechanics the wave properties of a particle are described by the de Broglie length, $\lambda = h/p$. One can introduce the average de Broglie length corresponding to the average value of a momentum, via $<\lambda>= h/<p>$ and rewrite Eq. (7.4) in the form,

$$\frac{\langle r\rangle}{\langle\lambda\rangle} >> 1. \tag{7.5}$$

This equation in fact provides the condition for the system to be classical or behave according to classical mechanics rules. However, to make this condition less obscure in physical terms, it is necessary to elucidate the role of temperature in it. Thus in spite of using $\lambda = h/p$, with the average value for momenta, we just take the thermal de Broglie wavelength as $<\lambda>$, in accordance to,

$$<\lambda>= \frac{h}{\sqrt{2\pi m k_B T}}. \tag{7.6}$$

Recall that $V/N = 4\pi <r>^3 /3$ and then $<r>$ is of the order of $n^{-1/3}$. Then, Eq. (7.5) takes the form,

$$\frac{2\pi m k_B T}{h^2 n^{-2/3}} >> 1. \tag{7.7}$$

We can define the parameter, T_{deg}, called as degeneracy temperature,

$$T_{deg} = \frac{h^2}{2\pi m k_B} n^{2/3}. \tag{7.8}$$

Then, the condition for the system to behave as classic is,

$$T/T_{deg} >> 1. \tag{7.9}$$

The classical approximation is better if the temperature of the system is high, or the system is less dense, or the particles are heavier. Otherwise, one should resort to quantum theory. We will see below, that classical approximation is adequate when $k_B T$ is much bigger than the energy determining distance between energy levels of a quantum system.

7.2 Elements of quantum mechanics

Concept of wave function. The properties of atomic objects in quantum mechanics are described by using an auxiliary function that is called as wave function (or state vector). The wave function describes the state of motion of a single particle. In general it is complex (not real), unique valued and continuous function of a coordinate, $\mathbf{r}$ and time t. The wave function, $\psi(\mathbf{r},t)$ satisfies certain differential equation that determines motion of a particle. It is called as Schrödinger equation and plays similar role in quantum mechanics as Newton equations of motion in classical mechanics. For example, the free motion of an electron with the determined momentum, p, is described by the wave function (corresponding to a plane de Broglie wave),

$$\psi(\mathbf{r},t) = A\exp[i(\mathbf{kr} - \omega t)], \tag{7.10}$$

where, $\mathbf{k} = \mathbf{p}/\hbar$ is the wave vector and $\omega = p^2/2m\hbar$ is the circular frequency.

The wave function depends on the coordinates and time. The number of coordinates equals to the number of degrees of freedom of the system. Denote a set of all independent coordinates at certain moment of time by the letter ξ. Defining ξ yields a point in an abstract space called as configuration space. The volume element in the configuration space is $d\xi$. For a system with one particle only, the configuration space coincides with Cartesian 3D space. In quantum mechanics the wave function is required to calculate

values of physical magnitudes in the state of motion determined by it. Restrict our attention to a single moment of time and omit time variable for the moment.

Specifically, it is assumed in quantum mechanics that,

$$\psi^*(\xi)\psi(\xi)d\xi = |\psi(\xi)|^2 d\xi, \tag{7.11}$$

is proportional to the probability, $W(\xi)$, that upon measuring the coordinates we will find them in the interval, ξ, $\xi + d\xi$. In many problems, the wave functions are or can be normalized,

$$\int |\psi(\xi)|^2 d\xi = 1. \tag{7.12}$$

The integration should be performed over all possible coordinate values. Then,

$$\rho(\xi) = |\psi(\xi)|^2 = \frac{dW(\xi)}{d\xi}, \tag{7.13}$$

is the probability density function of coordinates.

One example of the wave function that requires normalization is the one describing free motion of a particle, $\psi(\mathbf{r},t) = A\exp(i\mathbf{kr})$ with the momentum $\mathbf{p} = \hbar\mathbf{k}$. In order to perform normalization, we will define the wave function inside a big cubic volume (with the volume $V = L^3$) and apply boundary conditions at the borders of a volume. If the volume is quite big, the influence of boundary conditions on the motion of a particle will be negligible,

$$\psi(x,y,z) = \psi(x+L,y,z) = \psi(x,y+L,z) = \psi(x,y,z+L). \tag{7.14}$$

The boundary conditions will be satisfied, if the normalized wave functions are,

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{3/2}}\exp(i\mathbf{kr}),) \tag{7.15}$$

where

$$k_x = \frac{2\pi}{L}n_x, \quad k_x = \frac{2\pi}{L}n_x, \quad k_x = \frac{2\pi}{L}n_x, \tag{7.16}$$

with n_x, n_y, n_z are all positive and negative integers.

The introduction of boundary conditions leads to the discretization of the possible values of $\mathbf{k}$. In the limit $L \to \infty$ the distance between neighboring values of wave vector tends to zero and one recovers free motion of a particle in and infinite space.

The functions determined by Eq. (7.15) for all possible values of the wave vector form a set that satisfies the following conditions,

$$\int_\Omega d^3r\psi^*_{\mathbf{k}'}(\mathbf{r})\psi_{\mathbf{k}}(\mathbf{r}) = \delta_{\mathbf{k}'\mathbf{k}}, \tag{7.17}$$

where the Kronecker symbol in the r.h.s. is $\delta_{\mathbf{k}'\mathbf{k}} = \delta_{k'_x k_x}\delta_{k'_y k_y}\delta_{k'_z k_z}$, $\Omega = L^3$.

Moreover, these wave functions form a complete set such that any wave function, ψ, corresponding to an arbitrary state of motion of a particle in the volume Ω can be represented as a linear combination in terms of the basis functions, $\psi_{\mathbf{k}}(\mathbf{r})$,

$$\psi(\mathbf{r}) = \sum_{\mathbf{k}} a_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r}). \tag{7.18}$$

The coefficients of the expansion, $a_{\mathbf{k}}$, can be evaluated and interpreted straightforwardly. Namely,

$$\int_\Omega \psi(\mathbf{r})\psi^*_{\mathbf{k}'}(\mathbf{r})d^3r = \sum_{\mathbf{k}} a_{\mathbf{k}}\int_\Omega d^3r\psi_{\mathbf{k}}(\mathbf{r})\psi^*_{\mathbf{k}'}(\mathbf{r}) = \sum_{\mathbf{k}} a_{\mathbf{k}}\delta_{\mathbf{k}'\mathbf{k}} = a_{\mathbf{k}'}. \tag{7.19}$$

So, the result is,

$$a_{\mathbf{k}} = \int\limits_{\Omega} \psi(\mathbf{r})\psi^*_{\mathbf{k}}(\mathbf{r})d^3r. \tag{7.20}$$

If we assume that the functions $\psi(\mathbf{r})$ are normalized in the volume Ω, then the normalization condition leads to,

$$\int\limits_{\Omega} \psi^*(\mathbf{r})\psi(\mathbf{r})d^3r = 1 = \sum_{\mathbf{k}}\sum_{\mathbf{k}'} a_{\mathbf{k}}a^*_{\mathbf{k}'} \int\limits_{\Omega} d^3r\psi^*_{\mathbf{k}'}(\mathbf{r})\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{k}}\sum_{\mathbf{k}'} a_{\mathbf{k}}a^*_{\mathbf{k}'}\delta_{\mathbf{k}'\mathbf{k}} = \sum_{\mathbf{k}}|a_{\mathbf{k}}|^2. \tag{7.21}$$

Concerning the interpretation of the coefficients $a_{\mathbf{k}}$, we would like to note that according to Eq. (7.18), they express the contribution of the state with the momentum $\mathbf{p} = \hbar\mathbf{k}$ to the state determined by the wave function $\psi(\mathbf{r})$. The square of the absolute value, $|a_{\mathbf{k}}|^2$, describes the probability to find the value of momentum $\mathbf{p}$ for the system in the state described by $\psi(\mathbf{r})$. The last equation we derived, Eq. (7.21), in fact shows that the sum of probabilities of all possible values of momentum should equal to unity. We have discussed above that the wave function permits statistical interpretation. Therefore, it is necessary to explain how it (this concept) enters the calculations of any kind of average values (note that the notion of the operators of physical magnitudes has not been introduced so far).

7.3 Average values of a coordinate and momentum

According to our developments above, the probability density of a certain value of the vector $\mathbf{r}$ can be expressed through the normalized state function $\psi(\mathbf{r})$ as,

$$\rho = \psi^*(\mathbf{r})\psi(\mathbf{r}). \tag{7.22}$$

Apply then the theorem for mathematical expectation to write down the average (or mean) value $<\mathbf{r}>$ of the continuous random variable $\mathbf{r}$ as,

$$<\mathbf{r}> = \int \psi^*(\mathbf{r})\mathbf{r}\psi(\mathbf{r})d^3r. \tag{7.23}$$

The average refers to the state provided by $\psi(\mathbf{r})$. Similarly, the average of arbitrary function of the coordinate vector follows from,

$$<f(\mathbf{r})> = \int \psi^*(\mathbf{r})f(\mathbf{r})\psi(\mathbf{r})d^3r. \tag{7.24}$$

Evaluation of the mean value of momentum, $<\mathbf{p}> = \hbar <\mathbf{k}>$, in a given state is slightly more cumbersome. Recall the meaning of the coefficients $a_{\mathbf{k}}$ and the interpretation of $|a_{\mathbf{k}}|^2$. Thus, according to the general rule,

$$<\mathbf{p}> = \hbar\sum_{\mathbf{k}} a^*_{\mathbf{k}}\mathbf{k}a_{\mathbf{k}}. \tag{7.25}$$

We have the coefficients $a_{\mathbf{k}}$ defined by Eq. (7.20). Also, it is worth noting that,

$$\mathbf{k}\psi_{\mathbf{k}}(\mathbf{r}) = -i\nabla\psi_{\mathbf{k}}(\mathbf{r}), \tag{7.26}$$

as it follows from Eq. (7.15). Then, let us perform some mathematical transformations starting from Eq. (7.25). Namely, use expressions for $a^*_{\mathbf{k}}$ and for $a_{\mathbf{k}}$ and Eq. (7.26) to get,

$$<\mathbf{p}> = i\hbar\sum_{\mathbf{k}} \int d^3r'\psi^*(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}') \int d^3r\psi(\mathbf{r})\nabla\psi^*_{\mathbf{k}}(\mathbf{r}). \tag{7.27}$$

Recall, that the functions ψ and $\psi_{\mathbf{k}}$ have been normalized using cyclic boundary conditions, their respective values on the opposite boundaries of the cubic volume are equal. Therefore, if we perform integration by parts, it follows that,

$$\int d^3r\psi(\mathbf{r})\nabla\psi^*_{\mathbf{k}}(\mathbf{r}) = -\int d^3r\psi^*_{\mathbf{k}}(\mathbf{r})\nabla\psi(\mathbf{r}). \tag{7.28}$$

Thus,

$$< \mathbf{p} >= -i\hbar \int d^3r' \int d^3r \psi^*(\mathbf{r}') \left[\sum_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{r}') \psi^*_{\mathbf{k}}(\mathbf{r}) \right] \nabla \psi(\mathbf{r}). \tag{7.29}$$

The expression in the square brackets is just the delta-function of Dirac, i.e. $\sum_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{r}')\psi^*_{\mathbf{k}}(\mathbf{r}) = \delta(\mathbf{r}' - \mathbf{r})$. Consequently, the final form of our result for the mean value of momentum, in terms of the wave function determining the given state, is,

$$< \mathbf{p} >= \int d^3r \psi^*(\mathbf{r})(-i\hbar\nabla)\psi(\mathbf{r}). \tag{7.30}$$

This expression is valid for the general case of infinite space ($L \to \infty$) as well. By similar way of derivation it can be shown that for any rational function of momentum, the rule for calculation of the mean value is,

$$< F(\mathbf{p}) >= \int d^3r \psi^*(\mathbf{r}) F(-i\hbar\nabla)\psi(\mathbf{r}). \tag{7.31}$$

For example, the average kinetic energy of a particle in the state $\psi(\mathbf{r})$, is determined by the equation,

$$< \frac{p^2}{2m} >= \int d^3r \psi^*(\mathbf{r})(-\frac{\hbar^2\nabla^2}{2m})\psi(\mathbf{r}). \tag{7.32}$$

7.4 Operators for physical observables

Upon generalization of the results obtained above to get mean value for the coordinate and momentum in arbitrary state described by the normalized wave function, we state now that,

$$< F >= \int d^3r \psi^*(\mathbf{r}) \hat{F} \psi(\mathbf{r}), \tag{7.33}$$

where

$$\hat{F} = F_1(\mathbf{r}) + F_2(-i\hbar\nabla). \tag{7.34}$$

The operator $\hat{F}$ corresponds to a function $F = F_1(\mathbf{r}) + F_2(p)$. Generally, $\hat{F}$ is a differential operator. So, we wish to say that $\hat{F}$ corresponds to certain physical property or characteristics.

An operator is defined on a set of functions, if the law by which each function of a set is put in correspondence to another function from the same set. For example, the Laplace operator (denoted as ∇^2 or Δ) is a differential operator defined on the set of all twice differentiable functions. If an operator contains integration it is called as integral operator. A particular case of integral operator is given by a functional. It is an operator that while acting on any function from its defining set yields a constant. One example of the functional is the scalar product,

$$\langle \psi | \phi \rangle = \int \psi^*(\xi)\phi(\xi)d\xi. \tag{7.35}$$

Our desire now is to find a method of constructing the operators for physical magnitudes. Prior solving this problem it is worth to explore briefly what kind of conditions the operators of our interest should satisfy. First note that the action of an operator on the wave function to the right of it in Eq. (7.33) leads to the transformation of this function into a new one,

$$\hat{F}\psi = \psi'. \tag{7.36}$$

Such transformation should satisfy one of basic principles of quantum mechanics, namely the one called as superposition of states. Accordingly, it is required that the operators act as follows,

$$\begin{aligned}&\hat{F}(a\psi) = a\hat{F}\psi,\\ &\hat{F}(\psi_1+\psi_2) = \hat{F}\psi_1 + \hat{F}\psi_2.\end{aligned} \tag{7.37}$$

The operators satisfying such rules are called linear operators.

If the function F describes physical property, then its average value is real. This condition in accordance with Eq. (7.33) means that $< F >=< F >^*$ and the result can be expressed in terms of the integral equality,

$$\int d^3r\psi^*(\mathbf{r})\hat{F}\psi(\mathbf{r}) = \int d^3r\psi(\mathbf{r})\hat{F}^*\psi^*(\mathbf{r}). \tag{7.38}$$

This result is a particular case of a more general representation,

$$\int d^3r\psi^*(\mathbf{r})\hat{F}\phi(\mathbf{r}) = \int d^3r\phi(\mathbf{r})\hat{F}^*\psi^*(\mathbf{r}), \tag{7.39}$$

satisfied by the so-called self-adjoint or Hermitian operators. The abbreviated notation commonly used is, $\hat{F} = \hat{F}^\dagger$.

Comment in the bra and ket notation of quantum mechanics for the above: An operator is called Hermitian when it can always be flipped over to the other side, if it appears in a inner product, thus $< \phi|A\psi >=< A\phi|\psi >$ always, if A is Hermitian operator. The properties of inner products involving Hermitian operators are: $< \phi|A\psi >=< \psi|A\phi >^*$; $< \psi|A\psi >$ is real.

It is worth noting basic elements of the algebra of operators. The sum and difference of operators act as common $(\hat{A}+\hat{b})\psi = \hat{A}\psi + \hat{b}\psi$, $(\hat{A}-\hat{b})\psi = \hat{A}\psi - \hat{b}\psi$. The identity operator does nothing, $\hat{1}\psi = \psi$.

However, the product of operators $\hat{A}\hat{B}$ acts as consecutive action of each operator and depends on the order of application, in general $\hat{A}\hat{B}$ is not equal to $\hat{B}\hat{A}$. Therefore, it is convenient to define a commutator,

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{7.40}$$

If $\hat{A}$ and $\hat{B}$ happen to commute, then $[\hat{A},\hat{B}]\psi = 0$ for arbitrary function ψ, briefly this is expressed as $[\hat{A},\hat{B}] = 0$. Finally, to be comprehensive, mention the exponential of an operator $\exp\{\hat{A}\}$,

$$\exp\{\hat{A}\} = \hat{1} + \hat{A} + \frac{1}{2!}\hat{A}^2 + \frac{1}{3!}\hat{A}^3 + ... \tag{7.41}$$

that is given as power series.

The notation for simplest operators in quantum mechanics are collected in the table below.

Table 7.1: Examples of the operators of physical observables acting on the wave functions dependent of the coordinates of a particle.

Coordinate	Operators
Coordinate: $\mathbf{p}$; p_x, p_y, p_z.	$\hat{\mathbf{p}} = -i\hbar\nabla$; $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ $\hat{p}_y = -i\hbar\frac{\partial}{\partial y}$ $\hat{p}_z = -i\hbar\frac{\partial}{\partial z}$.
Kinetic energy: $K = \frac{\mathbf{p}^2}{2m} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}$.	$\hat{K} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2\nabla^2}{2m} = -\frac{\hbar^2}{2m}\Delta$ $= -\frac{\hbar^2}{2m}\left[\left(\frac{\partial^2}{\partial x^2}\right) + \left(\frac{\partial^2}{\partial y^2}\right) + \left(\frac{\partial^2}{\partial z^2}\right)\right]$.
Potential energy: $U(\mathbf{r},t) = U(x,y,z,t)$.	Operator of momenta $U(\mathbf{r},t) = U(x,y,z,t)$.
Moment: $\mathbf{L} = [\mathbf{r}\ \mathbf{p}] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix}$ $L_x = yp_z - zp_y$, $L_y = zp_x - xp_z$, $L_z = xp_y - yp_x$.	$\hat{\mathbf{L}} = [r\ \hat{\mathbf{p}}] = -i\hbar\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{vmatrix}$ $\hat{L}_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$, $\hat{L}_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$, $\hat{L}_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$.

7.5 Eigenfunctions and eigenvalues of the operators

As we have seen above Eq. (7.33) permits to obtain the average value of arbitrary property F in the state described by the wave function $\psi(\mathbf{r})$, if we know the operator $\hat{F}$ corresponding to this property. This rule,

$$< F >= \int d^3r\psi^*(\mathbf{r})\hat{F}\psi(\mathbf{r}), \tag{7.42}$$

can be applied not only for the average value, but to the mean square deviation of the property (in the state given by $\psi(\mathbf{r})$) as well. The deviation of the property from its average is, $\Delta F = F - < F >$. Then, the hermitian operator for the deviation is $\hat{(\Delta F)} = \hat{F} - < F >$. The mean square deviation, $< (\Delta F)^2 >$, follows from the equation,

$$< (\Delta F)^2 >= \int d^3r\psi^*(\mathbf{r})\hat{(\Delta F)}\hat{(\Delta F)}\psi(\mathbf{r}) = \int d^3r|\hat{(\Delta F)}\psi(\mathbf{r})|^2, \tag{7.43}$$

because $\hat{(\Delta F)}$ is hermitian. There should be some, unknown states in which the mean square deviation equals zero, i.e.

$$\int d^3r|\hat{(\Delta F)}\psi(\mathbf{r})|^2 = 0, \tag{7.44}$$

In other words, in such states the property F takes definite value. This is possible if,

$$\hat{(\Delta F)}\psi(\mathbf{r}) = 0, \tag{7.45}$$

or equivalently,

$$(\hat{F} - F)\psi(\mathbf{r}) = 0, \tag{7.46}$$

where we have put $< F >= F$ because the property in question takes definite value. The last equation is homogeneous, linear equation for the unknown function $\psi(\mathbf{r})$. We are interested in the solutions of this equation corresponding to non-zero, continuous, unique valued (uni-valued) functions $\psi(\mathbf{r})$ that would describe real states of physical systems.

In general, Eq. (7.46) has solutions only for certain values of the property F, these values can be interpreted as the parameters of Eq. (7.46) . Moreover, these parameters can form a discrete set, $F_1, F_2, ...$ or be continuous in the certain interval. These specific values are called as eigenvalues of the operator $\hat{F}$ (a set of eigenvalues determines spectrum), whereas the corresponding solutions of Eq. (7.46) are called as eigenfunctions of the operator $\hat{F}$. If the spectrum is discrete, we can use notation, $F_1, F_2, .., F_n, ..$ and ψ_{F_n}, and the integer numbers n are called as quantum numbers.

A very important property of the eigenvalues of hermitian operators is that they always take real values. The eigenvalues coincide with the average values of the corresponding physical observables in the states that are described by eigenfunctions of these operators. Because, the average values are real, the eigenvalues are real as well. In order to prove that eigenvalues of hermitian operators are real take Eq. (7.46) and multiply it by ψ^*, i.e. by complex conjugate of ψ,

$$\psi^* \hat{F} \psi - \psi^* F \psi = 0, \tag{7.47}$$

and rest the complex conjugate of this equation,

$$\psi \hat{F}^* \psi^* - \psi F^* \psi^* = 0. \tag{7.48}$$

From integration of the result we find,

$$(F - F^*) \int \psi^* \psi d^3 r = \int \psi^* \hat{F} \psi d^3 r - \int \psi \hat{F}^* \psi^* d^3 r. \tag{7.49}$$

However, the operator $\hat{F}$ is hermitian, see e.g. Eq. (7.39) , and therefore the eigenvalues are real, $F = F^*$.

7.6 On the properties of eigenfunctions of operators with discrete spectrum

Let the operator $\hat{F}$ has a discrete spectrum of the eigenvalues, F_n. Then the eigenfunctions of this operator satisfy the equation,

$$\hat{F} \psi_n = F_n \psi_n. \tag{7.50}$$

Write down the complex conjugate of this equation for the quantum number m,

$$\hat{F}^* \psi_m^* = F_m \psi_m^*. \tag{7.51}$$

Next, multiply from the left the first and the second equation by ψ_m^* and by ψ_n, respectively. Hence,nequation

$$\int d\xi (\psi_m^* \hat{F} \psi_n - \psi_n \hat{F}^* \psi_m^*) = F_n \int d\xi \psi_m^* \psi_n - F_m \int d\xi \psi_n \psi_m^* = (F_n - F_m) \int d\xi \psi_m^* \psi_n. \tag{7.52}$$

However, the operator $\hat{F}$ is hermitian and therefore,

$$(F_n - F_m) \int d\xi \psi_m^* \psi_n = 0. \tag{7.53}$$

If $m \neq n$, we observe that the eigenfunction corresponding to different eigenvalues are orthogonal,

$$\int d\xi \psi_m^* \psi_n = 0. \tag{7.54}$$

The physical meaning of the orthogonality of the eigenfunctions is that if we would measures an observable F, we will get F_n as the result in the state ψ_n and F_m in the state ψ_m. Let assume that normalization of eigenfunctions has been performed, then

$$\int d\xi \psi_m^* \psi_n = \delta_{mn}. \tag{7.55}$$

Another important property of the eigenfunctions for the operators with discrete spectrum, is that the set of eigenfunctions forms a complete set, i.e. any wave function (with the same boundary conditions) can be expanded using a complete set,

$$\psi(\xi) = \sum_n a_n \psi_n(\xi). \tag{7.56}$$

The summation should be performed over all values of the quantum number n. Then the coefficients of the expansion are straightforwardly obtained from,

$$a_n = \int d\xi \psi(\xi) \psi_n^*(\xi). \tag{7.57}$$

With all this knowledge, it is easy to express the mean value for the physical observable, $< F >= \int d\xi \phi^* F \psi$ in terms of the coefficients of the expansion above and the eigenvalues of the operator,

$$< F >= \sum_n F_n |a_n|^2. \tag{7.58}$$

Also, the normalization condition yields,

$$\int d\xi \psi^*(\xi) \psi(\xi) = 1 = \sum_n |a_n|^2. \tag{7.59}$$

In physical terms, the square of the absolute value of the coefficient a_n determines the probability that after measuring the observable F in the state determined by ψ we will obtain the value F_n.

7.7 Schrödinger equation

One of principal equations of quantum mechanics is the Schrödinger equation that determines evolution of states of quantum systems with time. It has the form,

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi, \tag{7.60}$$

where H is the Hamiltonian operator (it coincides with the operator of energy under absence of time dependence in it). Restricting to non-relativistic case, the operator $\hat{H}$ is the sum of the operators of kinetic and potential energy of a particle,

$$H = -\frac{\hbar^2}{2m} \nabla^2 + \hat{U}(r). \tag{7.61}$$

Our interest is in stationary states, thus assume that the Hamiltonian does no depend on time explicitly, $\partial H / \partial t = 0$. The solution of Schrödinger equation will be seeked for in the form,

$$\psi(\xi, t) = \psi(\xi) A(t). \tag{7.62}$$

Then, Eq. (7.60) is,

$$\frac{i\hbar \frac{\partial A(t)}{\partial t}}{A(t)} = \frac{H\psi(\xi)}{\psi(\xi)} = E, \tag{7.63}$$

where E is constant. This result implies two equations,

$$H\psi_E(\xi) = E\psi_E(\xi) \tag{7.64}$$

and

$$i\hbar\frac{\partial A(t)}{\partial t} = EA(t). \tag{7.65}$$

The Eq. (7.64) is the equation determining eigenvalues of the Hamiltonian operator, which is the energy operator under condition $\partial H/\partial t = 0$. The wave functions $\psi_E(\xi)$ correspond to states of the system in which the energy has definite values. The solution of Eq. (7.65) is given explicitly,

$$A(t) = \exp(-iEt/\hbar). \tag{7.66}$$

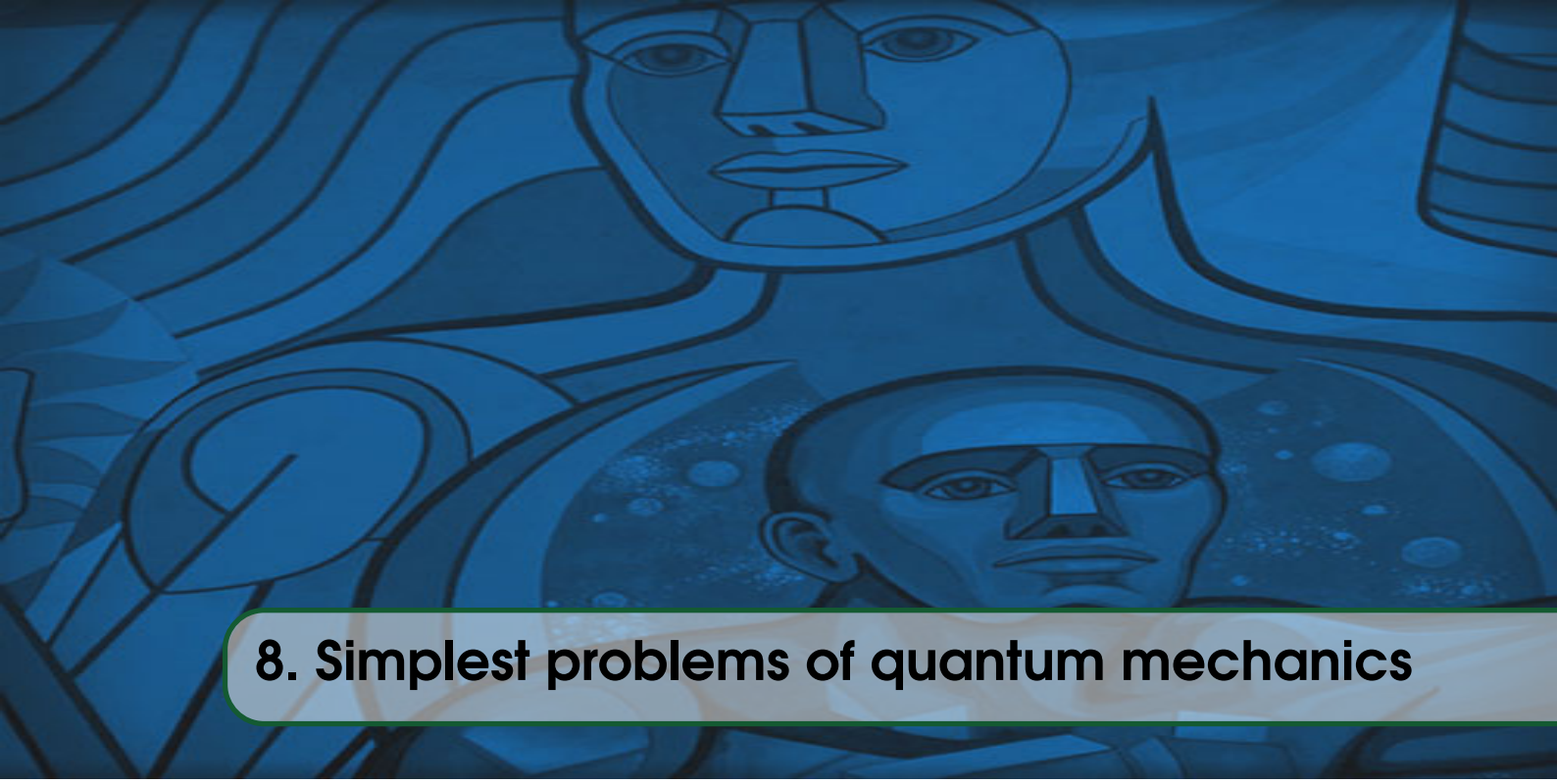
The states having definite energy are called as stationary states in quantum mechanics. The wave function describing stationary states then is,

$$\psi(\xi,t) = \psi(\xi)\exp(-iEt/\hbar) \tag{7.67}$$

The stationary states possess certain properties. For purposes of our text we just note that in such states, the average value of any physical magnitude (if its operator does not depend on time explicitly) is constant,

$$< F >= \int d\xi\psi^*(\xi,t)F\psi(\xi,t) = const. \tag{7.68}$$

Besides, the probability to find certain value of arbitrary physical magnitude in a stationary state does not depend on time.



8. Simplest problems of quantum mechanics

8.1 Particle in a one-dimensional potential square well

Consider a single particle free to move inside a one-dimensional box of the size L_x with impermeable walls, i.e. $U(0) = U(L) = \infty$; $U(x) = 0, 0 < x < L_x$. Here, $U(x)$ is the external field that determines the potential energy of the system.

The Schrödinger equation for stationary states is,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x). \tag{8.1}$$

The boundary conditions are,

$$\psi(x=0) = \psi(L_x) = 0. \tag{8.2}$$

General solution of the equation is,

$$\psi(x) = Csin(kx+\delta), \tag{8.3}$$

and

$$E = \frac{\hbar^2}{2m}k^2. \tag{8.4}$$

The unknown constants of the solutions, C, k, and δ are obtained from the boundary conditions and normalization condition,

$$\int_0^{L_x} |\psi(x)^2 dx = 1. \tag{8.5}$$

From the boundary conditions, we have $\delta = 0$ and $kL_x = \pi n$. Thus, we have

$$\psi(x) = Csin(kx), \qquad k = \pi n/L_x, \quad n = 1,2,3,... \tag{8.6}$$

The state with $n = 0$ is absent as it is equivalent to zero. From the normalization condition we obtain,

$$|C|^2 \int_0^{L_x} sin^2 kx dx = \frac{1}{2}|C|^2 \int_0^{L_x} (1 - cos2kx)dx = \frac{L_x}{2}|C|^2 = 1. \tag{8.7}$$

Then, $C = (2/L_x)^{1/2}$. So, the final result is,

$$\psi(x) = \sqrt{\frac{2}{L_x}} sin(\frac{\pi}{L_x} nx), \qquad E_n = \frac{\hbar^2}{2m}(\frac{\pi}{L_x})^2 n^2. \tag{8.8}$$

This exercise is a good illustration of the conclusions concerning the properties of eigenfunctions and eigenvalues of hermitian operators. The wave function of the ground state, $\psi_1(x)$ does not have nodes in the interval $0 < x < L_x$, it is real and positive. The next function, $\psi_2(x)$ describing the first excited state has one node at $x = L_x/2$, see the Fig. 8.1. This behavior is as expected to provide orthogonality,

$$\int_0^{L_x} \psi_1(x)\psi_2(x)dx = 0. \tag{8.9}$$

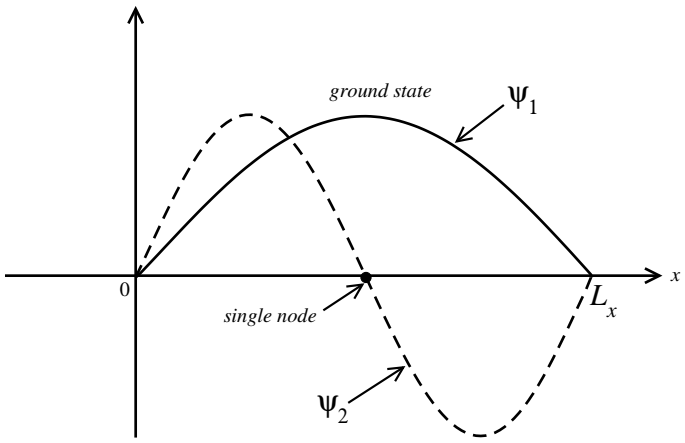


Fig. 8.1: Wave functions of the ground state, $\psi_1(x)$, and of the first excited state, $\psi_2(x)$ of a particle in one-dimensional box.

A set of eigenfunctions $\psi_n(x)$ forms a complete set of orthogonal and normalized functions,

$$\int_0^{L_x} \psi_n(x)\psi_{n'}(x)dx = \delta_{n,n'}. \tag{8.10}$$

This completeness permits to obtain some interesting relations.

■ **Example 8.1** Imagine that we have a wave function of the form, $\psi(x) = 1/\sqrt{L_x}$ defined in the interval, $0 < x < L_x$, and with boundary conditions, $\psi(0) = \psi(L_x) = 0$. Evidently, the function is normalized. Expand this function into series of eigenfunctions for a particle in the box,

$$\psi(x) = \sum_n C_n \psi_n(x), \tag{8.11}$$

where,

$$C_n = \int_0^{L_x} \psi_n(x)\psi(x)dx = \frac{\sqrt{2}}{L_x} \int_0^{L_x} sin\left(\frac{\pi}{L_x} nx\right) dx = \frac{\sqrt{2}}{\pi n}\left[1 - (-1)^n\right], \tag{8.12}$$

or explicitly,

$$C_n = \frac{2\sqrt{2}}{\pi n}, \quad n = 1,3,5,...; \quad C_n = 0, \quad n = 2,4,6,... \tag{8.13}$$

Now, substitute the result into Eq. (8.11) and next take $L_x = 2$,

$$\frac{1}{\sqrt{L_x}} = \frac{2\sqrt{2}}{\pi} \sum_{n=1,3,5,..} \frac{\sqrt{2}}{\sqrt{L_x}} \frac{1}{n} sin\left(\frac{\pi}{L_x} nx\right). \tag{8.14}$$

Thus,

$$\frac{\pi}{4} = \sum_{n=1,3,5,..} \frac{1}{n} sin\left(\frac{\pi}{2} nx\right). \tag{8.15}$$

! Not only that, but it is also surprising that a "square wave", which is horizontal for the uncountable number of x values in an entire interval, can be produced by adding only a countable number of sine functions, each of which is curved.

In addition, from the condition,

$$\sum_{n=1}^{\infty} |C_n|^2 = 1. \tag{8.16}$$

it follows, that,

$$\frac{\pi^2}{8} = \sum_{n=1,3,5,..} \frac{1}{n^2}, \tag{8.17}$$

from where it is easy to obtain,

$$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2}. \tag{8.18}$$

■

Exercise 8.1 Obtain this result by manipulation of the indices in the corresponding sums. ■

The solution of Schrödinger equation for a particle freely moving in a rectangular box with dimensions L_x, L_y and L_z ($V = L_x L_y L_z$) can be obtained applying similar development as in above. The energy levels then are,

$$E_{\mathbf{n}} = \frac{\hbar^2 \pi^2}{2m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right), \tag{8.19}$$

where n_x, n_y, n_z take integer values from 1 to ∞. The energy levels of a particle are determined by the triplets if indices n_x, n_y, n_z. In the particular case of a cubic box, the spacing between levels $n+1$ and n is of the order,

$$\Delta E_{n+1,n} = \frac{\hbar^2 \pi^2}{2m} \frac{2n+1}{V^{2/3}}. \tag{8.20}$$

This estimate, as obtained in the book of Davis (nitrogen molecule at 300K in the volume $1cm^3$), leads to the value $\approx 10^{-31} erg$, while $k_B T \propto 10^{-14} erg$. Thus the energy spacing for a free molecule in a macroscopic box is very small.

Our last comment in this section concerns statement of the problem for a system of N non-interacting particles in a box, say of quantum, ideal gas. The stationary Schrödinger equation for this case has the form,

$$\sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m_i}\right) \nabla_{r_i}^2 \psi = E\psi, \tag{8.21}$$

complemented by appropriate boundary conditions. The wave function of the entire system of particles is searched in the form,

$$\psi_{\mathbf{n}} = \prod_{i=1}^{N} \psi_{\mathbf{n}_i}, \tag{8.22}$$

where $\psi_{\mathbf{n}_i}$ refers to the wave function of a single particle. The energy levels of a system are characterized then by a set of quantum numbers, $n_{x_i}, n_{y_i}, n_{z_i}$. The expression for the eigenvalues of the operator can be obtained by generalizing the result given in Eq. (8.19), see e.g. book of Davis for explicit expressions [2].

8.2 Rigid rotor

Another exercise necessary for further developments and having pedagogical insight is the rigid rotator. A physical system in question is a point particle with mass m that is moving about an origin at 0 of the laboratory frame. The distance between the center and a particle is fixed at r_0. The problem has spherical symmetry with one coordinate fixed, however. The corresponding setup is shown in the Fig. 8.2.

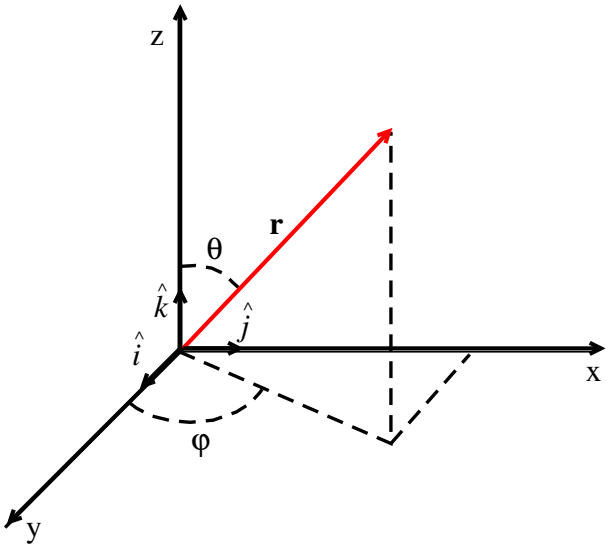


Fig. 8.2: Coordinate system for rigid rotor. The mass point constrained to move on a sphere of radius r_0. The **r** vector is written as $\mathbf{r} = r_0\left[\text{sin}(\theta)\text{cos}(\phi)\hat{i} + \text{sin}(\theta)\text{sen}(\phi)\hat{j} + \text{cos}(\theta)\hat{k}\right]$.

In order to solve the Schrödinger equation, first, we need to write down the Laplace operator in spherical coordinates and apply the constraint for the distance between a particle and the coordinate origin. It has the form,

$$\nabla_r^2\Big|_{r=r_0} = \frac{1}{r_0^2}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{8.23}$$

Then, the Schrödinger equation is,

$$-\frac{\hbar^2}{2I}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi(\theta,\phi) = E\psi(\theta,\phi), \tag{8.24}$$

where $I = mr_0^2$. Note that we have introduced the abbreviated notation, $\psi(\theta,\phi)$, rather than, $\psi(r_0,\theta,\phi)$. One just have in mind, that r_0 is the parameter of the problem and it will appear both in the eigenfunctions and in the eigenvalues.

Let us seek for the solution assuming that the wave function should have the following form, $\psi(\theta,\phi) = AY(\theta,\phi)$ where A is certain constant and $Y(\theta,\phi)$ is the function needed. It is convenient to rewrite Eq. (8.24) as follows,

$$\left[sin\theta\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right) + \frac{2IE}{\hbar^2}sin^2\theta\right]Y(\theta,\phi) = -\frac{\partial^2}{\partial\phi^2}Y(\theta,\phi). \tag{8.25}$$

Consequently, the desired function seems to have the form $Y(\theta,\phi) = \Theta(\theta)\Phi(\phi)$. With this assumption, Eq. (8.25) becomes transparent,

$$\frac{sin\theta}{\Theta(\theta)}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right)\Theta(\theta) + \beta sin^2\theta = \frac{1}{\Phi(\phi)}\frac{\partial^2}{\partial\phi^2}\Phi(\phi), \tag{8.26}$$

where $\beta = 2IE/\hbar^2$. Since the variables θ and ϕ are independent, each side of the equation above must be equal to a constant. Denote it as m^2. First solve simpler equation,

$$\frac{\partial^2}{\partial\phi^2}\Phi(\phi) = -m^2\Phi(\phi). \tag{8.27}$$

The solutions are,

$$\Phi(\phi) = A_m\exp(im\phi), \quad \Phi(\phi) = A_{-m}\exp(-im\phi). \tag{8.28}$$

The solution should be periodic according to the boundary condition,

$$\Phi(\phi) = \Phi(\phi + 2\pi). \tag{8.29}$$

This condition is satisfied if $m = 0, \pm1, \pm2, ...$ Thus,

$$\Phi(\phi) = A_m\exp(im\phi), \quad m = 0, \pm1, \pm2, .. \tag{8.30}$$

and m is called the magnetic quantum number. Still, normalization should be applied,

$$\int_0^{2\pi}\Phi(\phi)\Phi^*(\phi)d\phi = 1. \tag{8.31}$$

Then, the result is,

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}}\exp(im\phi), \quad m = 0, \pm1, \pm2, .. \tag{8.32}$$

Now, let us solve a more complicated equation for θ,

$$\frac{sin\theta}{\Theta(\theta)}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right)\Theta(\theta) + \beta sin^2\theta = m^2. \tag{8.33}$$

Introduce notation: $x = cos\theta, P(x) = \Theta(\theta)$. Note, that -1 $\leq x \leq$ 1. Then, Eq. (8.33) can be transformed into,

$$(1-x^2)\frac{d^2}{dx^2}P(x) - 2x\frac{d}{dx}P(x) + \left[\beta - \frac{m^2}{(1-x^2)}\right]P(x) = 0. \tag{8.34}$$

Exercise 8.2 Derive Eq. (8.34) from Eq. (8.33) by applying a spherical coordinate system. ■

It is the so-called Legendre differential equation for associated Legendre polynomials. This equation has nonzero solutions that are non-singular on [-1,1] only if $\beta = l(l+1)$ and l and m are integers with $0 \leq m \leq l$, or with trivially equivalent negative values. Thus, the eigenvalues are,

$$E = \frac{\hbar^2}{2I} l(l+1), \quad l = 0,1,2,... \tag{8.35}$$

Interestingly, the energy values do not depend on m, for each l there are $2l+1$ states with equal energy corresponding to different m. Such states are called degenerate states.

Still, to have a comprehensive expression for the eigenfunctions, one need to apply the normalization condition for the solution of Eq. (8.34). Specifically, if the the solutions of Eq. (8.34) are,

$$P_l^{|m|}(x) = P_l^{|m|}(cos\theta), \quad \Theta(\theta) = A_{lm} P_l^{|m|}(cos\theta), \tag{8.36}$$

$$A_{lm}^2 \int_0^{\pi} d\theta sin\theta [P_l^{|m|}(cos\theta)]^2 = 1. \tag{8.37}$$

The wave functions determining the states of rigid rotor,

$$Y_l^m(\theta,\phi) = \Theta_l^{|m|}(\theta)\Phi_m(\phi), \tag{8.38}$$

are called the spherical harmonics.

8.3 One-dimensional harmonic oscillator

In the classical mechanics, the harmonic oscillator is the system described by the Hamiltonian,

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} x^2. \tag{8.39}$$

The solutions of the equations of motion are well known,

$$x = x_0 sin(\omega t + \delta), \quad p = m\omega x_0 cos(\omega t + \delta), \tag{8.40}$$

where ω is the frequency of oscillations, x_0 is the amplitude and δ is the phase of oscillations. The energy, E, takes continuous values,

$$E = \frac{m\omega^2}{2} x_0^2. \tag{8.41}$$

In quantum mechanics, the momentum and the coordinate should be substituted by the operators. Then the Hamiltonian of the quantum oscillator is,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2} \hat{x}^2. \tag{8.42}$$

The task is to solve the Schrödinger equation, $\hat{H}\psi = E\psi$. In other words it is necessary to find the eigenfunctions and eigenvalues of the operator $\hat{H}$. In the coordinate space, apply the operators and the starting equation becomes,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + \frac{m\omega^2}{2} x^2 \psi(x) = E\psi(x), \tag{8.43}$$

with $-\infty < x < \infty$. Let us introduce the dimensionless length $\xi = x/l_0$, where $l_0 = \sqrt{\hbar/m\omega}$ is the characteristic length of the quantum oscillator called as the quantum amplitude. Thus, in the dimensionless units, the Schrödinger equation is,

$$-\frac{d^2}{d\xi^2}\psi(\xi) + \xi^2\psi(\xi) = \frac{2E}{\hbar\omega}\psi(\xi). \tag{8.44}$$

The asymptotic behavior of the wave function at $\xi \to \pm\infty$ can be established by considering the homogeneous counterpart of the differential equation written above,

$$-\frac{d^2}{d\xi^2}\psi(\xi) + \xi^2\psi(\xi) = 0. \tag{8.45}$$

The functional form of the solution in such limits is $\psi(\xi) \propto \exp(\pm\xi^2/2)$. However, using physical considerations (a particle should remain in space not far from the equilibrium distance) we discard the plus sign. Therefore, $\psi(\xi) \propto \exp(-\xi^2/2)$. So, the form of the solution would be,

$$\psi(\xi) = CH(\xi)\exp(-\xi^2/2), \tag{8.46}$$

where the unknown function $H(\xi)$ at infinity should decay faster than $\exp(-\xi^2/2)$. After substitution of Eq. (8.46) into (8.45), we obtain,

$$H^{''}(\xi) - 2\xi H^{'}(\xi) + \left(\frac{2E}{\hbar\omega} - 1\right) H(\xi) = 0. \tag{8.47}$$

It seems that the solution should be a polynomial (with even values of k to match the asymptotic behavior!), hence we assume that,

$$H(\xi) = \sum_{k\geq 0} a_k \xi^k. \tag{8.48}$$

By substitution, we obtain the equation,

$$\sum_{k\geq 2} a_k k(k-1)\xi^{k-2} - 2\sum_{k\geq 0} a_k k\xi^k + \left(\frac{2E}{\hbar\omega} - 1\right)\sum_{k\geq 0} a_k \xi^k = 0. \tag{8.49}$$

Then, after simple manipulation,

$$\sum_{k\geq 0}\left[a_{k+2}(k+2)(k+1) - 2ka_k + \left(\frac{2E}{\hbar\omega} - 1\right)a_k\right]\xi^k = 0. \tag{8.50}$$

The sum of the series equals to zero only, if each coefficient of the expansion vanishes. Thus, we obtain a recursion formula for the unknown coefficients a_k,

$$a_{k+2} = a_k\left[\frac{2k+1-2E/\hbar\omega}{(k+2)(k+1)}\right]. \tag{8.51}$$

This means that we need to specify both a_0 and a_1 to find all of the coefficients. In the limit of big k,

$$\lim_{k\to\infty} a_{k+2} \approx 2a_k/k. \tag{8.52}$$

Thus, $a_{2k} \approx 1/k!$ if k is even. In consequence the series

$$\sum_k \frac{\xi^{2k}}{k!} = \exp\left(\xi^2\right), \tag{8.53}$$

result in the function $H(\xi) \propto \exp(+\xi^2)$ growing faster than desirable, i.e. the wave function will not satisfy the boundary conditions. The only way out of this conundrum is that the series must be finite. Let us suppose that there exists certain value of n such that $a_{n+2} = 0$ whereas $a_n \neq 0$. This fact implies,

$$2n + 1 - \frac{2E}{\hbar\omega} = 0. \tag{8.54}$$

The resulting equation determines the energy levels $E = E_n$ for the harmonic oscillator,

$$E_n = \hbar\omega(n + 1/2), \quad n = 0, 1, 2, ... \tag{8.55}$$

that are evenly spaced or say equidistant along energy axis (see the Fig. 8.3).

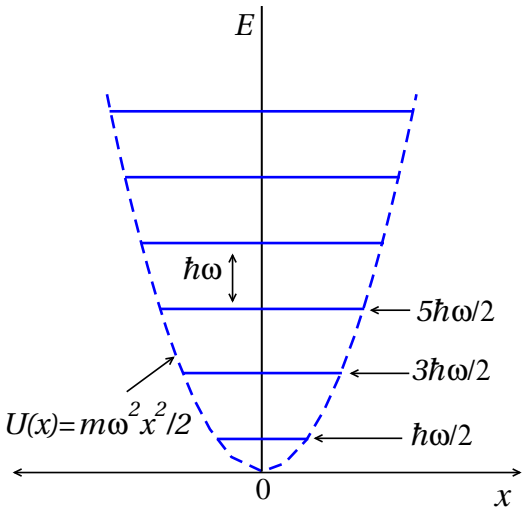


Fig. 8.3: Energy levels of quantum oscillator.

Comment. Note that imposing condition to truncate the series, one terminates either the odd series or the even series, because the recursion relation is spaced by two. One needs to separately impose the specific condition $a_k = 0$ for the other series (see Appendix I of Ref. [18]). This is fine because it can be shown that if the potential is symmetric, then the energy eigenfunctions can be taken to be either even or odd .

As a consequence of the developments above we have the following summarizing results,

$$a_{k+2} = a_k \frac{2(k-n)}{(k+2)(k+1)}, \tag{8.56}$$

and the function,

$$H(\xi) = H_n(\xi) = \sum_{k=0}^{n} a_k \xi^k, \tag{8.57}$$

that describes the Hermite polynomials under condition that the coefficient at a highest degree is chosen as $a_n = 2^n$. The explicit expressions for the Hermite polynomials follow from the definition,

$$H_n(\xi) = e^{\xi^2} \left(-\frac{d}{d\xi}\right)^n e^{-\xi^2} = e^{\xi^2/2} \left(\xi - \frac{d}{d\xi}\right)^n e^{-\xi^2/2}. \tag{8.58}$$

Returning to Eq. (8.46), we are able to write down the eigenfunctions of the oscillator as

$$\psi_n(\xi) = C_n H_n(\xi) \exp(-\xi^2/2), \tag{8.59}$$

the normalization constants follow from,

$$\int_{-\infty}^{\infty} \phi_n^2(x)dx = 1, \tag{8.60}$$

and equal to,

$$C_n = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}}. \tag{8.61}$$

The energy states are non-degenerate, one wave function corresponds to each value of energy. The ground state energy with $n = 0$ is the lowest energy,

$$E_0 = \frac{1}{2}\hbar\omega, \tag{8.62}$$

with non-zero energy. It is a consequence of the uncertainty principle of Heisenberg. The equations $< x^2 >= 0$ and $< p^2 >= 0$ cannot be satisfied simultaneously.

The wave function of the ground state,

$$\psi_0(\xi) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\xi^2/2}, \tag{8.63}$$

does not have nodes. However, the wave function of the first excited state ($E_1 = \frac{3}{2}\hbar\omega$),

$$\psi_0(\xi) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\xi^2/2}\sqrt{2}\xi, \tag{8.64}$$

has one node. See, the Fig. 8.4 for the sake of illustration.

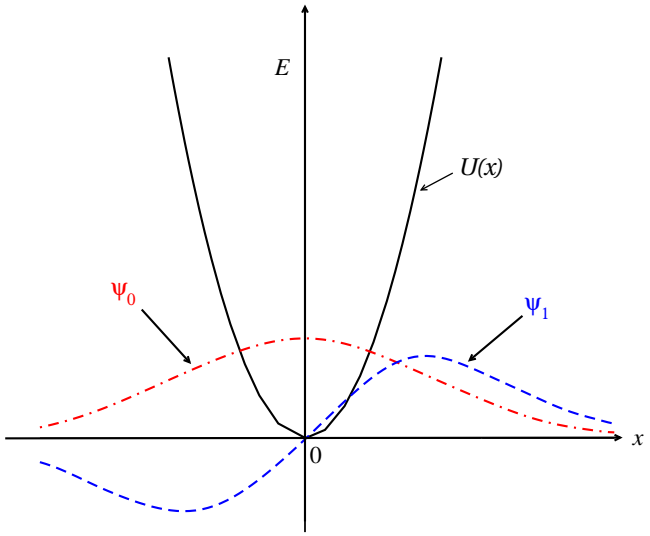


Fig. 8.4: Wave functions of the ground state and of the first excited state.

9. Diatomic molecule in a box

9.1 Classical versus quantum description

Let us consider a model for a diatomic molecule consisting of atoms, or more specifically of two mass points, m_1 and m_2 that are held together by the potential acting between $u(\mathbf{r})$. It is assumed that the vector $\mathbf{r}$ has origin at the coordinate of particle 1 and is pointed to particle 2, see the Fig. 9.1 for illustration. Consider a molecule in laboratory frame and point the vector $\mathbf{R}$ from the origin of the frame to the center of mass (COM) of the molecule.

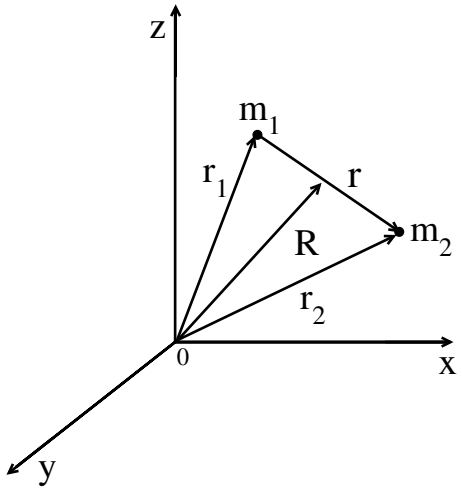


Fig. 9.1: Laboratory coordinate frame for a diatomic molecule.

The classical Hamiltonian of the model can be written as,

$$H_{cl} = \frac{1}{2}\sum_{i=1}^{2} m_i v_i^2 + u(r). \tag{9.1}$$

The model can be described by using the coordinates, R, r and angles θ and φ that determine orientation of the vector $\mathbf{r}$ in the COM frame (see the Fig. 9.1 for illustrative

purposes). The center of mass coordinate is defined as,

$$\mathbf{R} = \frac{1}{M}(m_1\mathbf{r}_1 + m_2\mathbf{r}_2), \tag{9.2}$$

where M is the mass of the molecule, $M = m_1 + m_2$. For the sake of convenience, let us introduce the reduced mass of a molecule as well, $m_{red} = m_1 m_2/(m_1 + m_2)$. In similar fashion as above, define the vector for the COM velocity,

$$\mathbf{v}_{COM} = \frac{1}{M}(m_1\mathbf{v}_1 + m_2\mathbf{v}_2). \tag{9.3}$$

The spherical polar coordinates of the COM frame can be used to describe vector of the diatomic molecule,

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 = r(sin\theta cos\varphi\mathbf{i} + sin\theta sin\varphi\mathbf{j} + cos\theta\mathbf{k}). \tag{9.4}$$

Then, from Eq. 9.4 the time evolution of the vector $\mathbf{r}$ (or in other words the relative velocity) is,

$$\mathbf{v}_r = \frac{d\mathbf{r}}{dt} = \mathbf{v}_2 - \mathbf{v}_1 = \dot{r}\hat{r} + r\boldsymbol{\omega}, \tag{9.5}$$

where in Eq. 9.5, $\hat{r}$ is unit vector along $\mathbf{r}$ and $\boldsymbol{\omega}$ is the angular velocity, $\boldsymbol{\omega} = \mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z$, with the components

$$\begin{aligned}
\omega_x &= \dot{\theta}cos\theta\cos\varphi - \dot{\phi}sin\theta sin\varphi, \\
\omega_y &= \dot{\theta}cos\theta\sin\varphi + \dot{\phi}sin\theta cos\varphi, \\
\omega_z &= -\dot{\theta}sin\theta.
\end{aligned} \tag{9.6}$$

With notation of Eq. 9.3 and 9.5, the kinetic energy expression in the initial Hamiltonian (see Eq. 9.1) can be written as follows,

$$\frac{1}{2}\sum_{i=1}^{2} m_i v_i^2 = \frac{1}{2}Mv_{COM}^2 + \frac{1}{2}m_{red}v_r^2. \tag{9.7}$$

Note that, $v_r^2 = \dot{r}^2 + r^2\omega^2$ from Eq. 9.5, because the unit vectors $\mathbf{r}$ and $\boldsymbol{\omega}$ are orthogonal. Then, the classical Hamiltonian from Eq. (9.1) is given by,

$$H_{cl} = \frac{1}{2}Mv_{COM}^2 + \frac{1}{2}I\omega^2 + \frac{1}{2}m_{red}\dot{r}^2 + u(r), \tag{9.8}$$

where $I = m_{red}r^2$. Each of these terms has straightforward physical meaning. Namely, the first one describes translational motion of a diatomic molecule, the second describes rotation of the molecule. Finally, the third term can be identified with the vibration of the bond length.

Prior to the application of quantum mechanics, let us make some simplifications. The first observation concerns the strength of the bond between atoms forming the diatomic molecules. We assume that the amplitude of vibrations of atoms is small, in other words the well of the potential keeping atoms together is quite deep and narrow. Thus, the expansion of the potential in Taylor series,

$$u(r) = u_0 + \frac{1}{2}\left(\frac{\partial^2 u}{\partial r^2}\right)\Big|_{r_0}(r - r_0)^2 + \cdots, \tag{9.9}$$

is truncated after the second term, $u_0 = u(r = r_0)$. Introduce notation, $x = r - r_0$ and $\lambda = \left(\frac{\partial^2 u}{\partial r^2}\right)\Big|_{r_0}$. Then, their vibrational classical Hamiltonian is,

$$H_{cl}^{vib} = \frac{1}{2}m_{red}\dot{x}^2 + \frac{1}{2}\lambda x^2 + u_0. \tag{9.10}$$

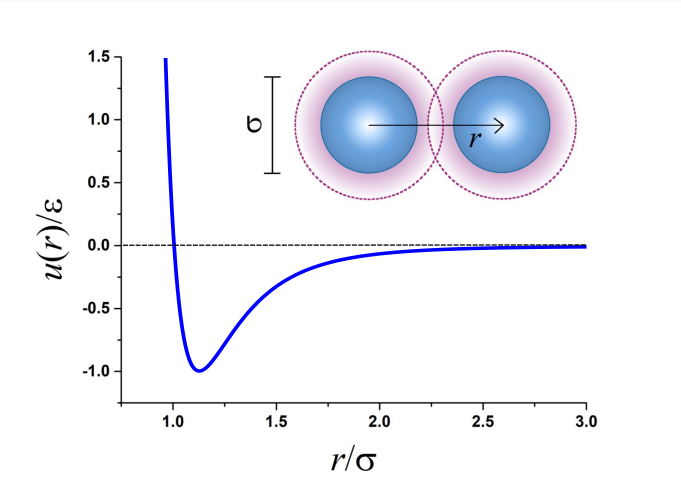


Fig. 9.2: The interaction potential holding atoms together in the molecule.

It is of the same form as we have explored in the previous chapter considering one-dimensional harmonic oscillator (with frequency $\omega = \sqrt{\lambda/m}$). One more approximation concerns the angular momentum I. For simplicity it is assumed that rotation of the molecule does not depend on the bond length. It is quite reasonable approximation, at least at high temperatures where classical mechanics is valid. Therefore, we assume that $I = m_{red}r^2 \approx m_{red}r_0^2 = I_0$. With these approximations three terms of the classical Hamiltonian, $H_{cl}^{tr}, H_{cl}^{rot}$ and H_{cl}^{vib} are completely independent of one another.

Now, we turn our attention to the quantum mechanical description of the problem as more general and more precise. If the velocities in Eq. (9.1) are substituted by momenta, apply necessary expressions for the operators and write down the starting form of the Schrödinger equation as,

$$-\sum_{i=1}^{2}\frac{\hbar^2}{2m_i}\nabla_{r_i}^2\psi + u(|\mathbf{r}_1 - \mathbf{r}_2|)\psi = E\psi. \tag{9.11}$$

However, it can be written in the form that takes into account symmetry of the problem by using the COM frame, i.e. in $\mathbf{R}$, r, θ and φ and the corresponding form of the operators,

$$(H^{tr} + H^{rot} + H^{vib})\psi = E\psi, \tag{9.12}$$

where for translational motion,

$$H^{tr}\psi = -\frac{\hbar^2}{2M}\nabla_{\mathbf{R}}^2\psi = E\psi. \tag{9.13}$$

Next, for rotational motion of a molecule we have,

$$H^{rot}\psi = -\frac{\hbar^2}{2I_0}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right]. \tag{9.14}$$

Finally,

$$H^{vib}\psi = -\frac{\hbar^2}{2I_0}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \left[\frac{\lambda}{2}(r - r_0)^2 + u_0\right]\psi. \tag{9.15}$$

Note that, the first term here, in Eq. (9.15), differs from what we have studied in the previous chapter concerning one-dimensional harmonic oscillator. Specifically, the first

term in Eq. 9.15 is the three dimensional radial part of the Laplace operator written in spherical coordinates. Moreover, the approximation for I has been applied to yield I_0.

The solution of the problem is performed in close similarity and along lines studied in the previous chapter. It yields the eigenfunctions,

$$\psi_{\mathbf{n},lm,\nu} = \psi^{tr}_{\mathbf{n}} \psi^{rot}_{lm} \psi^{vib}_{\nu}. \tag{9.16}$$

The first two functions on the r.h.s. of this equation are familiar. However, the eigenfunction corresponding to the vibration operator is,

$$\psi^{vib}_{\nu} = \frac{A_\nu}{r} H_\nu \left(\frac{r - r_0}{x_0} \right) \exp\left\{ \frac{-(r-r_0)^2}{2x_0^2} \right\}, \tag{9.17}$$

where $x_0 = (\hbar^2/m_{red}\lambda)^{1/4}$ and H_ν are the Hermite polynomials.

The eigenvalues of the total Hamiltonian are,

$$E_{\mathbf{n},lm,\nu} = E^{tr}_{\mathbf{n}} + E^{rot}_{lm} + E^{vib}_{\nu} + u_0, \tag{9.18}$$

where $u_0 = u(r = r_0)$ is a constant, and

$$E^{rot}_{lm} = \frac{\hbar^2}{2I_0} l(l+1), \quad E^{vib}_{\nu} = \hbar\omega(\nu + 1/2). \tag{9.19}$$

Exercise 9.1 Obtain the expression for the Laplace operator in spherical coordinates,

$$\mathbf{\Delta} = \nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta \frac{\partial\psi}{\partial\theta} \right) + \frac{1}{sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2} \right], \tag{9.20}$$

starting from Cartesian variables.

■

9.2 Hydrogen atom in a box

The final part of this chapter concerns intrinsically similar problem. It can be viewed as an application of the model for diatomic molecule studied above. Consider a hydrogen atom consisting of a proton and an electron. In contrast to the development above, the Coulomb interaction between the electron and the proton cannot be approximated by using the concept of harmonic oscillator.

To begin with, let us use the result from the **homework** of the previous section. We return to the Eqs.(9.11)-(9.15) without applying approximation $I = I_0$ and without expansion for the potential energy,

$$-\frac{\hbar^2}{2M}\nabla^2_{\mathbf{R}}\psi - \frac{\hbar^2}{2m_{red}r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r} \right) - \frac{\hbar^2}{2m_{red}r^2}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial\psi}{\partial\theta} \right) + \frac{1}{sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2} \right] + u(r)\psi = E\psi. \tag{9.21}$$

Note that, $M = m_p + m_e \approx m_p$, $m_{red} = m_p m_e/(m_p + m_e) \approx m_e$ and the COM coordinate, $\mathbf{R} \approx \mathbf{r}_p$. Our focus is not in the motion of an atom as total, it is in fact trivial. Rather we are interested in the relative motion with respect to the COM.

Thus, we would like to solve the Schrödinger equation of the form,

$$-\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r} \right) + \frac{1}{2mr^2}\hat{L}^2 + u(r)\psi = E\psi. \tag{9.22}$$

Note, that the expression in the square bracket in Eq. (9.21) acting on ψ, is the squared angular momentum operator $\hat{L}^2$.

Because the radial and angular variables are decoupled, we look for the solution of the Schrödinger equation in the form of a product,

$$\psi(\mathbf{r}) = R(r)Y_{lm}(\theta,\varphi). \tag{9.23}$$

The function $R(r)$ is called the radial function, it is determined by the equation,

$$-\frac{\hbar^2}{2mr}\frac{\partial^2}{\partial r^2}(rR(r)) + \frac{\hbar^2 l(l+1)}{2mr^2}R(r) + u(r)R(r) = ER(r). \tag{9.24}$$

Our experience and the form of the equation permit to choose the solution as,

$$\chi(r) = rR(r). \tag{9.25}$$

The unknown function $\chi(r)$ satisfies the equation,

$$\left\{-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + U_l(r)\right\}\chi(r) = E\chi(r), \tag{9.26}$$

where we introduced the notation,

$$U_l(r) = u(r) + \frac{\hbar^2 l(l+1)}{2mr^2}. \tag{9.27}$$

The function $U_l(r)$ plays role of the effective potential energy. The first term is the Coulomb attraction between the proton and electron, $-e^2/r$, whereas the second term describes repulsion and prohibits falling the electron on the attractive center. The variable r is defined in the interval $0 \leq r < \infty$.

Explore first, how the function $\chi(r)$ should behave at small and large values of the distance. If $r \to 0$, we preserve only the two dominating terms in the Eq. (9.26),

$$-\frac{\hbar^2}{2m}\frac{d^2}{dr^2}\chi(r) + \frac{\hbar^2 l(l+1)}{2mr^2}\chi(r) = 0. \tag{9.28}$$

It has been assumed that $ru(r) \to 0$. Then, the solution is taken in the form $\chi = const \times r^k$. We compute the first and second derivatives,

$$\frac{d\chi(r)}{dr} = const \times kr^{(k-1)}, \qquad \frac{d^2\chi(r)}{dr^2} = const \times k(k-1)r^{(k-2)}, \tag{9.29}$$

where substituting these into Eq. (9.28), we obtain,

$$-k(k-1)r^{(k-2)} + l(l+1)r^{(k-2)} = 0, \tag{9.30}$$

and finally,

$$-k(k-1) + l(l+1) = 0, \tag{9.31}$$

and then either $k = -l$ or $k = l+1$. The first solution is non-physical because then the radial function, $R(r) = \chi(r)/r$ would tend to ∞, if $r \to 0$, i.e. the particle will "fall" on the attractive center. Thus, we leave with $k = l+1$, and,

$$\chi(r) = const \times r^{l+1}, \quad r \to 0. \tag{9.32}$$

Next, consider the limit $r \to \infty$. Preserving solely the dominating terms, leads to the equation,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dr^2}\chi(r) = E\chi(r). \tag{9.33}$$

The solution should have the form,

$$\chi(r) \propto e^{\alpha r}, \quad r \to \infty. \tag{9.34}$$

Substitution of Eq. (9.34) into (9.33) determines the unknown constant,

$$\alpha = \pm\sqrt{-\frac{2mE}{\hbar^2}}. \tag{9.35}$$

For $E > 0$, we will have motion with continuous values of energy; note that α is imaginary, i.e.,

$$\alpha = \pm i\sqrt{\frac{2mE}{\hbar^2}}. \tag{9.36}$$

Consequently, the wave function is oscillatory and the sign determines direction of the spherical wave, either propagating from the center or to the center.

The bound states of the system are determined by $E < 0$. Then, in order to satisfy condition that $R(r) \to 0$ at $r \to \infty$, we will preserve only one value,

$$\alpha = -\sqrt{\frac{2m|E|}{\hbar^2}}, \tag{9.37}$$

that leads to the wave function,

$$\chi(r) \propto \exp\left(-\sqrt{\frac{2m|E|}{\hbar^2}}r\right). \tag{9.38}$$

Finally, for bound states, taking into account the behavior of the function χ in two limiting cases, we can write down the radial wave function as follows,

$$R(r) = r^l \exp\left(-r\sqrt{-\frac{2mE}{\hbar^2}}\right) w(r). \tag{9.39}$$

This expression provides correct behavior of the radial wave function in two limiting cases, whereas the function $w(r)$ is written formally to mimic the behavior of $R(r)$ at intermediate distances, evidently the form of $w(r)$ depends on the specific shape of the potential $u(r)$.

Our deliberations starting from Eq. (9.24) have a quite general character. Therefore, redo some principal issues focusing in the case of a hydrogen atom. As it has been mentioned above, in this specific case, $u(r) = -e^2/r$. Moreover it is convenient to introduce a dimensionless variable $\rho = r/a$, where a should be a certain characteristic length for the specific problem. With this variable, the radial Schrödinger Eq. (9.26) is,

$$\left\{-\frac{\hbar^2}{2ma^2}\frac{d^2}{d\rho^2} + \frac{\hbar^2 l(l+1)}{2ma^2\rho^2} - \frac{e^2}{a}\frac{1}{\rho}\right\}\chi(r) = E\chi(r) \tag{9.40}$$

or

$$\left[-\frac{d^2}{d\rho^2} + \frac{l(l+1)}{\rho^2} - \frac{2mae^2}{\hbar^2}\frac{1}{\rho}\right]\chi(r) = \frac{E}{(\hbar^2/2ma^2)}\chi(r). \tag{9.41}$$

Let us choose the length scale by the relations $mae^2/\hbar^2 = 1$, then $a = a_B = \hbar^2/me^2$ is the so-called Bohr radius. This assumption fixes the energy scale of the problem as well,

$$\frac{\hbar^2}{2ma^2} = \frac{me^4}{2\hbar^2}, \tag{9.42}$$

the r.h.s. determines energy in Rydberg units. With these definitions,

$$\left[-\frac{d^2}{d\rho^2}+\frac{l(l+1)}{\rho^2}-\frac{2}{\rho}\right]\chi=\varepsilon\chi, \tag{9.43}$$

where $\varepsilon=-E/(\hbar^2/2ma^2)=-E/(me^4/2\hbar^2)$, the sign minus is introduced because $E<0$, thus ε is positive! General form of the wave function follows from Eq. (9.39),

$$\chi(\rho)=\rho^{l+1}e^{-\rho\sqrt{\varepsilon}}w(\rho). \tag{9.44}$$

Then, the differential equation for the unknown function $w(\rho)$ is,

$$xw^{''}(x)+[2(l+1)-x]w^{'}(x)+[1/\sqrt{\varepsilon}-(l+1)]w(x)=0, \tag{9.45}$$

where $x=2\rho\sqrt{\varepsilon}$. The solution is searched for in the form of the polynomial for the variable x. It is lengthy and the detailed procedure can be found in the textbooks of quantum mechanics. Here, we would like just to mention that the steps of this deliberation have similarity with the solution of the problem for one-dimensional oscillator in chapters 7 and 8. We need only to note that the result is,

$$w(x)=\sum_{k}^{n-l-1}a_kx^k. \tag{9.46}$$

Truncation of the polynomial, note the upper value in the summation above, to satisfy limiting behavior of the wave function leads to the quantization of energy, such that,

$$\varepsilon=\frac{1}{n^2},\quad E_n=-\frac{me^4}{2\hbar^2}\frac{1}{n^2},\quad n=1,2,3,... \tag{9.47}$$

where n is the principal quantum number, while the orbital quantum number, l, is delimited by n, l takes values $0,1,..,n-1$. To be more comprehensive at this stage, we just write down the explicit resulting expression for the radial wave function of the problem,

$$R_{n,l}(r)=-C_{nl}\left(\frac{2\rho}{n}\right)^l\mathrm{e}^{-\rho/n}L_{n+l}^{2l+1}\left(\frac{2\rho}{n}\right), \tag{9.48}$$

where,

$$C_{nl}=\sqrt{\frac{4(n-l-1))!}{n^4a_B^3[(n+l)!]^3}}, \tag{9.49}$$

and $L_q^p(x)$ are the associated Laguerre polynomials. By this we practically finish the solution of the problem concerned with the motion of an electron in the field of Coulomb potential for bounded states, sometimes it is called as Kepler problem.

The total wave function of the problem, according to Eq. (9.23), then is,

$$\psi_{n,l,m}(\mathbf{r})=Y_{l,m}(\theta,\varphi)R_{n,l}(r). \tag{9.50}$$

It depends on three quantum numbers: n - the principal quantum number, l - orbital quantum number and m that is called as the magnetic quantum number (it determines the energy levels in magnetic field. The energy, E_n depends on the principal quantum number only. The fact that it does not depend on m is due to the symmetry of the Hamiltonian relative to rotation around arbitrary axis in space. On the other hand, the fact that E_n does not depend on the orbital quantum number l is the manifestation of additional symmetry of the Hamiltonian that yields one additional integral of motion. This issue can be learned in the quantum mechanics textbooks. Here, we would like just to emphasize that the radial functions, $R_{n,l}(r)$ determine the probability density $4\pi r^2R_{n,l}^2(r)$ of the distribution

for the "electron cloud" along the radius r. For example, in the case of the ground state the probability density is,

$$r^2 R_{1,0}^2 = \frac{4r^2}{a_B^3} e^{-2r/a_B}, \tag{9.51}$$

it reaches maximum value at $r = a_B$. In other words, the most probable distance at which one can find an electron in the ground state equals to the Bohr radius. The functions describing the excited state have several maxima, however.

10. Canonical ensemble

10.1 Postulates of the theory of ensembles

The theory of statistical ensembles is constructed by using a set of postulates.

Definition: An ensemble is a set of systems consisting of M copies of a single (original) system at equilibrium. The number of copies M is very big, actually at some stage, we will perform a limit $M \to \infty$. The original system as well as all its copies is considered under the same thermodynamic conditions, we denominate them as constraints. Thermodynamic conditions are described by a small number of parameters, e.g. N, V and T. The name for each ensemble is given in accordance with thermodynamic constraints. Namely, the microcanonical ensemble is a set of isolated systems under the conditions that N, V and E are chosen fixed. Here, E denotes the total energy of the system. An ensemble of systems with N, V and T fixed is called as canonical ensemble. We will consider the isobaric-isothermal ensemble defined by the N, P and T parameters. And finally, the grand canonical ensemble is defined by μ, V and T conditions, where μ is the chemical potential. This latter ensemble is open, as it permits interchange of particles between the reservoir and the system.

In above, we have used notation N for the number of particles, in other words restricting to a system of single species. This has been assumed just for simplicity. In fact, all the definitions and machinery are applicable for any multicomponent equilibrium system. Then one should define $\mathbf{N}$ describing composition of the system in terms of the number of particles of each species $N_1, N_2, ...$ for microcanonical, canonical and isobaric ensemble. In this situation, the grand canonical ensemble should be defined by using thermodynamic constraints for each species, $\boldsymbol{\mu}$ with $\mu_1, \mu_2, ..$ describing chemical potential for each species.

A system subject to certain constraints can generally evolve in time through a huge number of quantum states (or microstates) as the effect of intra- and inter-molecular interactions. We do not have device to explore the system in different microstates. Thus, it is assumed that the number of copies of the original system is so big, that all possible (dependent on constraints) microstates are present in the ensemble. Even more, the number of copies describing each unknown microstate α, denote it as M_α, is big. If we refer to the microstate as the random event, then the probability, p_α, of finding the system in the microstate α (or say to find the microstate α) is,

$$p_\alpha = \frac{M_\alpha}{M}. \tag{10.1}$$

If we would know the probability distribution for an ensemble, then the ensemble average of a certain mechanical property F could be obtained using,

$$< F >= \sum_{\alpha} p_{\alpha} F_{\alpha}. \tag{10.2}$$

Frequently, the theory of ensembles is developed starting from a premise (or postulate) that the thermodynamic or long-time average value of the mechanical variable is equal to the ensemble average value of this variable.

Is it true? Rigorously speaking the answer is no, because within mental construction an ensemble contains all possible microstates conformed with external constraints, whereas in the time evolution the system would not "visit" all possible microstates even if the observation time is quite large. In order to proved that the postulate is valid, one would need to show that practically only the most probable microstates contribute to the average. In other words, it is necessary to demonstrate that the series $p_{\alpha}F_{\alpha}$, if ordered by the magnitude of the values of p_{α} is rapidly converging.

According to the explanation given by T. Davis in Ref. [2], the justification of this postulate is known as the ergodic hypothesis, it applies to molecular systems governed by classical or quantum mechanics. Although the hypothesis has been rigorously proved under certain restrictive conditions, in general, it remains a postulate. Its validity is largely accepted due to the extensive empirical evidence supporting the agreement between theoretical predictions and experimental observations. A more detailed discussion of the ergodic hypothesis at fundamental level one can find in the monograph by R. Balescu [19].

The second postulate of ensemble theory is that in an ensemble representative of an isolated thermodynamic system, the systems of the ensemble are distributed "uniformly" or say with equal probability (or frequency) over all possible quantum states of the system. This is known as the postulate of equal a priori probability. If this statement is explained in terms of time, one would say that over a long period of time the fraction of time an isolated system spends in any one quantum state (consistent with N, V and E constraints) is the same as the fraction of time it spends in any other distinct quantum state.

The validity of these postulates can be questioned, but the results of ensemble theory never contradict experimental findings. On the other hand, ensemble theory possesses predictive power guiding experimental research in various situations.

10.2 Towards entropy

It is of interest to comment how the concept of entropy enters the theory of ensembles. Recall how the ensemble is constructed and restrict to an isolated system with given parameters, N, V and energy E. First, we pick up a single system with such constraints. Next, add the second, the third and following copies of the system all subject to the same values of external parameters. We are unaware, or better say, we are unable to describe the microstates in which each of the systems is. Therefore, just pertinent comment is that the systems are in different microstates permitted by the external parameters. So the states are undefined or chaotically distributed according to the values of external thermodynamic parameters. Let us use the concept of absence of knowledge about the distribution of microstates and define a property that characterizes disorder or absence of knowledge about microstates distribution.

Let us define the entropy as the measure of uncertainty of a microstate of the system from the ensemble. The entropy should be a number, thus it is a functional of the probability distribution of microstates. In order to find how the microstates are distributed, recall that the equilibrium state in thermodynamics is unique and corresponds to the maximum of entropy, this observation permits to search the law according to which the microstates are distributed.

Thus, we are looking for the functional of entropy such that: a) entropy should be zero if there is no absence of knowledge, i.e. say if one of microstates is certain and all other are impossible. Introduce notation, p_α, as the probability of microstate α. If we have $p_\alpha = 1$ and $p_{\alpha'} = 0$ for all $\alpha' \neq \alpha$, then we know everything about the microstates distribution, i.e. there is no uncertainty.

The second condition is the following. At a given number of microstates, the entropy should be at maximum if all the microstates are equally probable. Moreover, the entropy should increase, if the number of equally probable microstates increases. In this situation our absence of knowledge, or uncertainty is at maximum. Finally, the entropy must be additive, or in other words the entropy of the system composed of several independent parts should be the sum of the entropy's of the parts.

Let us show that all these conditions are satisfied if the entropy is defined as,

$$S = -k_B \sum_{\{\alpha\}} p_\alpha \ln p_\alpha, \tag{10.3}$$

where k_B is the Boltzmann constant.

Consider now how the requirements described above match the definition of entropy. According to the first condition above, assume that all the systems of an ensemble are in the same (certain) state α_0 such that $p_{\alpha_0} = 1$. Then, the probability to find any other state for the systems in an ensemble $\alpha \neq \alpha_0$ is zero. Thus,

$$S = -k_B p_{\alpha_0} \ln p_{\alpha_0} - k_B \sum_{\alpha \neq \alpha_0} p_\alpha \ln p_\alpha. \tag{10.4}$$

Here, we will take into account that $\lim_{x\to 0} x \ln x = 0$. Hence,

$$S = -k_B 1 \times \ln 1 - k_B \sum_{\alpha \neq \alpha_0} 0 \times \ln 0 = 0. \tag{10.5}$$

Consequently, the entropy of an ensemble, in which uncertainty of the microstate is absent, equals to zero. Now, let us show that entropy is at maximum, if all the microstates are equiprobable. Search for the extremum condition in the definition of entropy,

$$\delta S = -k_B \sum_{\alpha=1}^{G} [\delta p_\alpha \ln p_\alpha + p_\alpha \delta(\ln p_\alpha)] = -k_B \sum_{\alpha=1}^{G} [\ln p_\alpha + 1] \delta p_\alpha = 0, \tag{10.6}$$

where G is the total number of microstates for the systems forming ensemble. It is evident, that if all the variations δp_α would be independent, then the extremum of entropy is determined by equations to put all the coefficients $\ln p_\alpha + 1$ equal to zero. However, because the probabilities are normalized,

$$\sum_{\{\alpha\}} p_\alpha = 1, \tag{10.7}$$

then,

$$\sum_{\alpha=1}^{G} \delta p_\alpha = 0, \tag{10.8}$$

and therefore only $G - 1$ variations of all G are independent. Consequently, it is more convenient to use the method of Lagrange multipliers and search the unconditional extremum of the functional,

$$L = S - \lambda \sum_{\{\alpha\}} p_\alpha, \tag{10.9}$$

where λ is the Lagrange multiplier that should be determined afterwards. Now, the desired extremum follows from the equation,

$$\delta L = -k_B \sum_{\alpha=1}^{G} [\ln p_\alpha + 1 + \lambda/k_B]\delta p_\alpha = 0. \tag{10.10}$$

Now, all the variations are independent and the extremum is determined by the condition,

$$\ln p_\alpha + 1 + \lambda/k_B = 0, \tag{10.11}$$

or

$$p_\alpha = \exp[-\lambda/k_B - 1]. \tag{10.12}$$

It is clear that $p_\alpha = const$, because λ and k_B are constants. Using the normalization condition for the probability of the microstates, Eq. (10.7), we obtain,

$$p_\alpha = \frac{1}{G}, \tag{10.13}$$

and the entropy from the definition Eq. (10.3) attains the final form,

$$S = k_B \ln G. \tag{10.14}$$

Thus, the entropy has a maximum, moreover the entropy increases with increasing the number of microstates. Finally, let us show that S is additive. Let the system consists of two independent subsystems. The probability of a microstate of the system as total is the product of probabilities of certain microstates for each subsystem,

$$p_{\alpha_{12}} = p_{\alpha_1} p_{\alpha_2}. \tag{10.15}$$

Then, according to the definition of entropy,

$$S = -k_B \sum_{\alpha_{12}} p_{\alpha_{12}} \ln p_{\alpha_{12}} = -k_B \sum_{\{\alpha_1,\alpha_2\}} p_{\alpha_1} p_{\alpha_2} \ln p_{\alpha_1} - k_B \sum_{\{\alpha_1,\alpha_2\}} p_{\alpha_1} p_{\alpha_2} \ln p_{\alpha_2}. \tag{10.16}$$

Because of normalization of probabilities of microstates for each subsystem, we obtain,

$$S = S_1 + S_2. \tag{10.17}$$

To conclude, we observed that the entropy according to the definition in Eq. (10.3) really is a measure of uncertainty of the microstate of the system in an ensemble. Thus, the distribution of probability of the microstate can be searched for from the extremum of entropy at certain additional constraints. The distribution should be such that the ensemble averages would lead to certain values of physical properties and these averages are different for different ensembles.

Let us now proceed to the systematic procedure to find the probability distribution of microstates in different ensembles. Recall that the only useful knowledge and starting point is the postulate of the microcanonical ensemble of equal a priori probability.

10.3 Statement of the problem: Canonical ensemble

To begin with, consider a large set of systems (its number is M) each of volume V and each with number of particles N. All these systems have been equilibrated at temperature T. Each of the systems has been restricted by diathermal walls to provide equilibration from starting conditions, such that heat can pass through the walls but the amount of matter in the system is fixed. We can consider the entire set of M systems as a single large

super-system. The original systems are the subsystems of the super-system. Let us assume that the super-system is isolated from the surrounding by adiabatic walls.

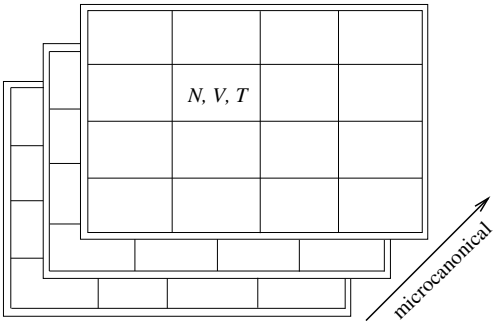


Fig. 10.1: Schematic representation of the super-system in question. It is composed of M subsystems, each with volume V, each with N particles and in equilibrium at temperature T. The super-system is isolated by adiabatic walls.

Recalling the applications of quantum mechanics discussed in previous chapters, we state that a subsystem can be in quantum states with discrete energies E_i, $(i = 1,2,3,...)$ which depend on V and N. Therefore, the super-system that we deal with, can be described by the set of numbers, $\nu_1,\nu_2,\nu_3,...$, where ν_i is the number of subsystems in the i-th quantum state. According to the setup,

$$\sum_i \nu_i = M. \tag{10.18}$$

In addition, because the super-system is isolated, we have,

$$\sum_i \nu_i E_i = E_{tot}, \tag{10.19}$$

where E_{tot} is the total energy. It is a certain constant.

A given set of numbers, $\nu_1,\nu_2,\nu_3,...$ is denominated as distribution. Then, the super-system in question for a specific distribution of these numbers can be in any of

$$\Omega(\{\nu\}) = \frac{M!}{\prod_i(\nu_i!)} \tag{10.20}$$

distinct states. So we have the number of states for a given distribution.

The total number of distinct states possible for the super-system is,

$$\Omega_T = \sum_{\{\nu_1,\nu_2,...\}} \Omega(\{\nu\}), \tag{10.21}$$

where the summation is over all distribution sets of the numbers $(\nu_1,\nu_2,...)$ that are consistent with the imposed constraints for the number of subsystems M and for the total energy of isolated super-system E_{tot}.

Specify now our objective. We are looking for the probability p_i that a system with volume V and the number of particles N, in equilibrium at temperature T, is in the quantum state (microstate) with E_i. At a given set of numbers $\nu_1,\nu_2,...$, or equivalently with the assumed distribution set, the fraction of systems (belonging to a single super-system) in such state is,

$$\nu_i/M. \tag{10.22}$$

On the other hand, if we take a super-system and construct the microcanonical ensemble of super-systems, the postulate of equal a priori probability applies. According to this

postulate, the probability that a super-system picked at random from the microcanonical ensemble is in the microstate with the distribution set defined by $\nu_1, \nu_2, ...$ is,

$$\Omega(\{\nu\})/\Omega_T. \tag{10.23}$$

With this probability determined, we are able to say that the average fraction of subsystems in the state E_i is,

$$\sum_{\nu_1,\nu_2,...} \nu_i \frac{\Omega(\{\nu\})}{\Omega_T} =< \frac{\nu_i}{M} > . \tag{10.24}$$

The right hand side term of this equation would become the probability that a subsystem picked at random from a canonical ensemble (N, V, T) is in the state E_i, if the limit $M \to \infty$ is performed. So, we have a recipe to find the desired probability by using the l.h.s. as,

$$p_i = \lim_{M\to\infty} \sum_{\nu_1,\nu_2,...} \nu_i \frac{\Omega(\{\nu\})}{\Omega_T}. \tag{10.25}$$

Apparently, it looks as quite difficult task. However, the calculation can be performed because for large values of $\Omega(\{\nu\})$ it is dominated by its largest value, i.e. has a form like delta-function. Let us demonstrate this feature by considering a particular simple case.

Assume that there are only two quantum states, or microstates, for the subsystem. Then, the expression for $\Omega(\{\nu\})$ given by Eq. (10.20) is,

$$\Omega(\{\nu\}) = \frac{M!}{\nu_1!(M-\nu_1)!}. \tag{10.26}$$

We would like to show that Ω attains maximum at $\nu_1 = M/2$. If the numbers involved in the calculations are big, resort to the Stirling approximation, $\nu! \approx (\nu/e)^\nu$. Then the previous Eq. (10.26) becomes,

$$\Omega(\{\nu\}) = \frac{M^M}{\nu_1^{\nu_1}[(M-\nu_1)^{(M-\nu_1)}]} \tag{10.27}$$

If the maximum value at $\nu_1 = M/2$ exists, it is convenient to work with the deviation from this maximum, introduce the variable $m = \nu_1 - M/2$. Evaluate the value of $\Omega(\{\nu\})$ at $\nu_1 = M/2$ and denote it as Ω_{max}. Then, the ratio, $\Omega(\{\nu\})/\Omega_{max}$ is,

$$\Omega(\{\nu\})/\Omega_{max} = \frac{[(1-\frac{2m}{M})^{(m-M/2)}]}{[(1+\frac{2m}{M})^{(m+M/2]}} \tag{10.28}$$

To simplify the result, let us apply the following representation of the exponent,

$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n. \tag{10.29}$$

Then,

$$\Omega(\{\nu\})/\Omega_{max} = \frac{[(1-\frac{2m}{M})^m]}{[(1+\frac{2m}{M})^m]}, \tag{10.30}$$

under the condition that M is big. It is convenient to rescale the variable m as follows, $m = \xi M/2$, where ξ is the positive fraction. Then, we observe,

$$\lim_{M\to\infty} \Omega(\{\nu\})/\Omega_{max} = \lim_{M\to\infty} \left(\frac{1-\xi}{1+\xi}\right)^{\xi M/2} = 0, \quad if \quad \xi > 0. \tag{10.31}$$

Otherwise the l.h.s. equals to unity if $\xi = 0$.

As conclusion, we see that in the expression defining the probability by Eq. (10.25), only the terms yielding maximum of $\Omega(\{\nu\})/\Omega_{tot}$ matter. Above, in the illuminating simple exercise, we assumed the distribution set. With this knowledge, we are able to apply the approximation to evaluate $\Omega(\{\nu\})$ in Eq. (10.25) to obtain,

$$p_i = \lim_{M\to\infty} \frac{\nu_i^* \Omega_{max}}{M\Omega_{max}} = \lim_{M\to\infty} \frac{\nu_i^*}{M}, \tag{10.32}$$

where $\nu_1^*, \nu_2^*, ...$ is the distribution set of numbers maximizing the value for $\Omega(\{\nu\})$. Thus, to find the desired distribution, we must look for the maximum of $\Omega(\{\nu\})$ subject to the constraints for the numbers $\nu_1, \nu_2, ...$ given by Eqs. (10.18) and (10.19).

The distribution that maximizes $\Omega(\{\nu\})$ is the same that maximizes the functional $\ln\Omega(\{\nu\})$. The latter procedure offers certain advantages. So, we need to solve the equations,

$$\sum_i \left(\frac{\partial \ln\Omega}{\partial \nu_i}\right)\Bigg|_{\{\nu^*\}} \delta\nu_i = 0. \tag{10.33}$$

Additional constraints follow from Eqs. (10.18) and (10.19),

$$\sum_i \delta\nu_i = 0; \quad \sum_i E_i \delta\nu_i = 0. \tag{10.34}$$

As in the previous subsection, we apply the method of Lagrange multipliers to make the variations $\delta\nu_i$ independent. Thus, we obtain,

$$\sum_i \left[\left(\frac{\partial ln\Omega}{\partial \nu_i}\right)\Bigg|_{\{\nu^*\}} - \alpha - \beta E_i\right] \delta\nu_i = 0, \tag{10.35}$$

where α and β are the undetermined multipliers. Using the definition of Ω from Eq. (10.20) and taking the logarithm, we have,

$$\begin{aligned} \ln(\Omega(\{\nu\})) &= \ln(M!) - \ln\left(\prod_i \nu_i!\right) = \ln(M!) - \sum_i \ln(\nu_i!) \\ &= M\ln M - \sum_i \nu_i \ln\nu_i, \end{aligned} \tag{10.36}$$

where Stirling approximation for large numbers was used. Then take derivative by any of the numbers ν to get,

$$\begin{aligned} \left(\frac{\partial \ln\Omega}{\partial \nu_i}\right) &= \frac{\partial}{\partial \nu_i}\{(\nu_1+\nu_2+\cdots)\ln(\nu_1+\nu_2+\cdots) - (\nu_1\ln\nu_1 + \nu_2\ln\nu_2 + \cdots)\} \\ &= \ln M - \ln\nu_i. \end{aligned} \tag{10.37}$$

Therefore, the result for the Eq. (10.35) is,

$$\ln\left(\frac{M}{\nu_i}\right) - \alpha - \beta E_i = 0, \tag{10.38}$$

for any index i. Consequently, the probability of finding the system from canonical ensemble in the state E_i is,

$$p_i = \lim_{M\to\infty}\left(\frac{\nu_i^*}{M}\right) = e^{-\alpha-\beta E_i}. \tag{10.39}$$

Still, this is not the final form of the result. We should use the normalization condition for the probability to eliminate the multiplier α. Hence,

$$\sum_i p_i = 1 = \sum_i e^{-\alpha-\beta E_i}, \tag{10.40}$$

with the result,

$$\exp(-\alpha) = 1/\sum_i e^{-\beta E_i} = 1/Q_N. \tag{10.41}$$

Now, the final form of the result for the probability is,

$$p_i = \frac{1}{Q_N} e^{-\beta E_i}, \tag{10.42}$$

where $Q_N = \sum_i e^{-\beta E_i}$ is called the partition function or the sum of states of the canonical ensemble. We will see slightly later that Q_N contains all the information concerning thermodynamic properties of the system described by canonical ensemble.

10.4 Thermodynamic properties from partition function

Recall that we have the expression for the probability of finding the microstate of the system described by canonical ensemble,

$$p_i = \frac{1}{Q_N} e^{-\beta E_i}; \qquad Q_N = \sum_i e^{-\beta E_i}, \tag{10.43}$$

where Q_N is the partition function. The internal energy of the system under N, V and T conditions is,

$$U = \sum_i p_i E_i = \frac{1}{Q_N} \sum_i E_i e^{-\beta E_i}, \tag{10.44}$$

where E_i is the energy of the system in the i-th quantum state. In addition, let us define the pressure as,

$$P = \sum_i p_i P_i = \frac{1}{Q_N} \sum_i P_i e^{-\beta E_i}, \tag{10.45}$$

where,

$$P_i = -\left(\frac{\partial E_i}{\partial V}\right)_N. \tag{10.46}$$

This latter equation requires explanation. Specifically, the reversible work that has to be done **on** a closed system (with N fixed) in the energetic state i to increase its volume by dV is $dE_i = (\partial E_i/\partial V)_N dV = -P_i dV$. With these premises, we observe that the thermodynamic properties of the system are hidden in the partition function. Indeed, Eqs. (10.44) and (10.45) can be written as,

$$U = -\left(\frac{\partial \ln Q_N}{\partial \beta}\right)_{N,V}, \tag{10.47}$$

and

$$P = \frac{1}{\beta}\left(\frac{\partial \ln Q_N}{\partial V}\right)_{N,\beta}. \tag{10.48}$$

As it was mentioned at the end of the previous chapter, the Lagrange multiplier β is not determined so far. Our intention is to evaluate it. To do that, it seems reasonable to calculate the partition function for a system with well defined thermodynamic properties. Therefore, consider an ideal gas of mono-atomic particles and assume that the particles are

distinguishable to use the result coming from classical mechanics treatment. The energy of a given quantum state of such gas restricted to a cubic volume is,

$$E = \sum_{j=1}^{N} \frac{\pi^2\hbar^2}{2m_j V^{2/3}}(n_{x_j}^2 + n_{y_j}^2 + n_{z_j}^2), \tag{10.49}$$

where the quantum numbers $n_{x_1}, n_{y_1}, n_{z_1}, ..., n_{z_N}$ is a certain set of positive integers. According to the definition of the partition function Q_N, we need to perform summation over all possible quantum states, i.e. over all possible set of integer $3N$-tuples. Let us do that, starting from the definition of Q_N,

$$Q_N = \sum_{n_{x_1}=1}^{\infty} ... \sum_{n_{z_N}=1}^{\infty} \exp\left[-\beta \sum_{j=1}^{N} \frac{\pi^2\hbar^2}{2m_j V^{2/3}}(n_{x_j}^2 + n_{y_j}^2 + n_{z_j}^2)\right] = \prod_{j=1}^{N} q_j, \tag{10.50}$$

where,

$$q_j = \sum_{n_x=1}^{\infty}\sum_{n_y=1}^{\infty}\sum_{n_z=1}^{\infty} \exp\left[-\beta \frac{\pi^2\hbar^2}{2m_j V^{2/3}}(n_x^2 + n_y^2 + n_z^2)\right], \tag{10.51}$$

is the one-particle partition function. It is not straightforward to perform summation here. Therefore, we would like to substitute summation by integration. At the first step, let us formally introduce the grid in terms of the variables $\Delta n_x = \Delta n_y = \Delta n_z = 1$. Then,

$$q_j = \sum_{n_x=1}^{\infty}\sum_{n_y=1}^{\infty}\sum_{n_z=1}^{\infty} \Delta n_x \Delta n_y \Delta n_z \exp\left[-\beta \frac{\pi^2\hbar^2}{2m_j V^{2/3}}(n_x^2 + n_y^2 + n_z^2)\right]. \tag{10.52}$$

Next, let us absorb β and V into the variables,

$$\xi = \frac{\beta^{1/2}}{V^{1/3}} n_x, \quad \zeta = \frac{\beta^{1/2}}{V^{1/3}} n_y, \quad \eta = \frac{\beta^{1/2}}{V^{1/3}} n_z. \tag{10.53}$$

With these notation the one-particle partition function becomes,

$$q_j = \frac{V}{\beta^{3/2}} \sum_{\xi}^{\infty}\sum_{\zeta}^{\infty}\sum_{\eta}^{\infty} \Delta\xi \Delta\zeta \Delta\eta \exp\left[-\beta \frac{\pi^2\hbar^2}{2m_j}(\xi^2 + \zeta^2 + \eta^2)\right], \tag{10.54}$$

where each of the summations starts from $\beta^{1/2}/V^{1/3}$ and terminates at infinity with the step $\beta^{1/2}/V^{1/3}$. The step of summation can be made very small by assuming that the volume of the system is sufficiently large. Then, the summation can be substituted by integration as follows,

$$\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} d\xi d\zeta d\eta \exp\left[-\beta \frac{\pi^2\hbar^2}{2m_j}(\xi^2 + \zeta^2 + \eta^2)\right] = \frac{1}{8}\left(\frac{2\pi m_j}{\pi^2\hbar^2}\right)^{3/2}. \tag{10.55}$$

In consequence, the partition function of the N-particle system is,

$$Q_N = \frac{V^N}{\beta^{3N/2}} \prod_{j=1}^{N} \left(\frac{\pi m_j}{2\pi^2\hbar^2}\right)^{3/2}, \tag{10.56}$$

and the equation of state can be immediately obtained from Eq. (10.48), $P = N/(\beta V)$. By comparison with the equation of state already known for an ideal gas, we arrive at the result,

$$\beta = \frac{1}{k_B T}. \tag{10.57}$$

The internal energy then is,

$$U = \frac{3}{2} N k_B T. \tag{10.58}$$

In fact we have thermodynamics incomplete because entropy is missing. To do that, consider isothermal expansion for a closed, i.e. NVT system,

$$dU = TdS - PdV. \tag{10.59}$$

However, according to the definition of internal energy, $U = \sum_i p_i E_i$, we get,

$$dU = \sum_i E_i dp_i + \sum_i p_i dE_i = -k_B T d(\sum_i p_i \ln p_i) - PdV. \tag{10.60}$$

Exercise 10.1 Show this equation explicitly. ■

By comparison of Eq. (10.59) and Eq. (10.60) we can identify the differential of entropy as,

$$dS = -k_B d(\sum_i p_i \ln p_i). \tag{10.61}$$

According to thermodynamics dS is an exact differential, therefore the r.h.s. of Eq. (10.61) is also an exact differential. Thus we are able to integrate Eq. (10.59) to get,

$$S = -k_B \sum_i p_i \ln p_i + C = \frac{U}{T} + k_B \ln Q_N + C, \tag{10.62}$$

where C is the integration constant. One can choose $C = 0$ without loss of generality. Then,

$$S = \frac{U}{T} + k_B \ln Q_N. \tag{10.63}$$

This result should be compared with thermodynamic expression defining the Helmholtz free energy, A,

$$S = \frac{U}{T} - \frac{A}{T}, \tag{10.64}$$

to yield,

$$A = -k_B T \ln Q_N, \tag{10.65}$$

that represents fundamental relation for the canonical ensemble thermodynamics. To summarize, thermodynamics of the system described by canonical ensemble follows from,

$$P = k_B T \left(\frac{\partial \ln Q_N}{\partial V} \right)_{N,T}; \quad S = k_B \left(\frac{\partial (T \ln Q_N)}{\partial T} \right)_{N,V}; \quad \mu = -k_B T \left(\frac{\partial \ln Q_N}{\partial N} \right)_{V,T}, \tag{10.66}$$

where μ is the chemical potential.

10.5 Energy fluctuations

In the NVT system described by canonical ensemble the energy is not constant or say it fluctuates around it mean value. We would like to evaluate the magnitude of these fluctuations. Let us define the mean square deviation of energy in the canonical ensemble,

$$\sigma_E^2 = \overline{(E - <E>)^2}, \tag{10.67}$$

where E denotes the energy of the system (U can be used as well), two equivalent notations, the overline and $< ... >$, are used for averaging over ensemble. Recall that,

$$p_i = \frac{1}{Q_N(V,T)} \exp[-\beta E_i(N,V)]. \tag{10.68}$$

So,

$$\sigma_E^2 = \bar{E^2} - 2\overline{E<E>} + \overline{<E>^2} =< E^2 > -<E>^2 = \sum_i p_i E_i^2 - < E >^2 . \tag{10.69}$$

Let us evaluate the average of the squared energy first,

$$\begin{aligned} \frac{1}{Q_N} \sum_i E_i^2 \exp(-\beta E_i) &= -\frac{1}{Q_N} \frac{\partial}{\partial \beta} \left(\sum_i E_i e^{-\beta E_i} \right) = -\frac{1}{Q_N} \frac{\partial}{\partial \beta} (\bar{E} Q_N) \\ &= -\frac{1}{Q_N} \left[Q_N \frac{\partial \bar{E}}{\partial \beta} \right] - \bar{E} \frac{\partial ln Q_N}{\partial \beta} = k_B T^2 \frac{\partial \bar{E}}{\partial T} + < E >^2 . \end{aligned} \tag{10.70}$$

Thus,

$$< E^2 > - < E >^2 = k_B T^2 \frac{\partial < E >}{\partial T}. \tag{10.71}$$

If $< E >$ is the thermodynamic (internal) energy, we obtain,

$$\sigma_E^2 = k_B T^2 C_V, \tag{10.72}$$

where C_V is the heat capacity at constant volume. It is of interest to estimate the magnitude of the energy fluctuations with respect to the mean energy. In the case of ideal gas we will have,

$$\frac{\sigma_E}{<E>} = \frac{(k_B T^2 C_V)^{1/2}}{<E>} \propto O\left(\frac{1}{\sqrt{N}} \right). \tag{10.73}$$

11. Grand Canonical Ensemble

11.1 Statement of the problem: grand canonical ensemble

The grand canonical ensemble is a large collection (M is the large number, at the end of the procedure the limit of M to infinity will be performed) of copies of the original system that has volume V. The original system has been thermally equilibrated with a thermal bath at temperature T. Moreover, the original system has been chemically equilibrated with a bath at chemical potential μ [1]. In other words, we permit mass transfer between the bath and the system or say interchange of particles between them. The bath should contain similar particles as the system. We restrict attention to a single-component system for simplicity. Extension for mixtures can be performed straightforwardly, however.

One example of the physical system we are thinking of in using grand canonical ensemble is the fluid in porous media or in a single pore. The fluid in the pore is in equilibrium with the reservoir of the same particles. The chemical potential for the fluid in the pore and in the reservoir are equal at equilibrium. However, the density of the fluid in the reservoir (in the bulk phase) is different from the fluid density in the pore.

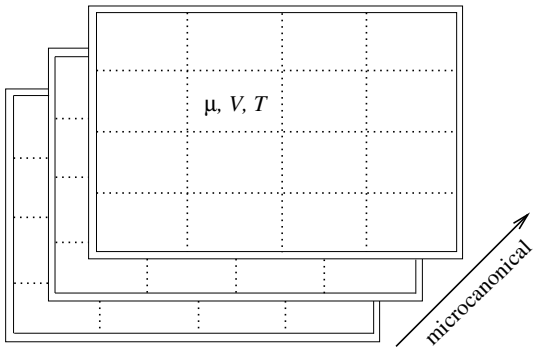


Fig. 11.1: Schematic representation of the grand canonical ensemble.

The systems of a grand canonical ensemble (i.e., copies of the original system) are open, each of them has a fixed volume, but the boundaries are permeable to both heat and matter exchange with a reservoir. The number of particles in the systems is determined by

[1] The chemical potential is the rate of change of the free energy at constant temperature and volume.

the chemical potential value, the number of particles is not fixed, it fluctuates around the mean value. Let us form a single super-system from the systems described by a grand canonical ensemble (see the figure) and isolate it to take advantage of the equal apriori probability in what follows.

In a given state of the super-system we will have a distribution ν, consisting of numbers $\nu_i(N)$. These numbers describe how many systems are in the state i with the number of particles N. So, we have the numbers like that: $\nu_1(0)$, then $\nu_1(1), \nu_2(1), \nu_3(1), ...$, then $\nu_1(2), \nu_2(2), \nu_3(2), ...$ and so on $\nu_1(3), \nu_2(3), \nu_3(3), ...$ In other words, we have the systems in different energy states with different occupation by particles. This occupation number can be thought as the internal quantum number for the state i as well.

In order to proceed, it is straightforward to apply similar reasoning as we have done for canonical ensemble in the previous chapter. To start, the number of possible states of the single super-system is,

$$\Omega(\{\nu\}) = \frac{M!}{\prod\limits_{i,N} \nu_i(N)!}. \tag{11.1}$$

Next, the constraints, similar to the procedure for the canonical ensemble, in the present case are,

$$M = \sum_i \sum_N \nu_i(N), \tag{11.2}$$

and

$$E_{tot} = \sum_i \sum_N \nu_i(N) E_i(N, V), \tag{11.3}$$

for the total energy of the isolated super-system. However, in the present case of the grand canonical ensemble we have one more, additional constraint,

$$N_{tot} = \sum_i \sum_N \nu_i(N) N. \tag{11.4}$$

All these ingredients above are necessary to search for the distribution of microstates in the grand canonical ensemble. Apply very similar arguments as in the previous chapter concerned with the canonical ensemble to arrive at the equation almost equal to Eq. (10.35),

$$\sum_{i,N} \left(\frac{\partial \ln \Omega(\{\nu\})}{\partial \nu_i(N)} \right) \Bigg|_{\{\nu^*\}} \delta\nu_i(N) = 0, \tag{11.5}$$

subject to the following constraints, which follow from Eqs. (11.2) and (11.3),

$$\sum_i \sum_N \delta\nu_i(N) = 0; \qquad \sum_i \sum_N E_i \delta\nu_i(N) = 0; \qquad \sum_i \sum_N N \delta\nu_i(N) = 0, \tag{11.6}$$

To make the variations $\delta\nu_i$ independent, we apply the method of Lagrange multipliers, α, β, and γ. This leads to,

$$\sum_{i,N} \left[\left(\frac{\partial \ln \Omega(\{\nu\})}{\partial \nu_i(N)} \right) \Bigg|_{\{\nu^*\}} - \alpha - \beta E_i - \gamma N \right] \delta\nu_i(N) = 0, \tag{11.7}$$

where γ is an additional Lagrange multiplier because we have one more constraint given by Eq. (11.4) here. The terms in square bracket in Eq. (11.7) should be zero because the variations $\delta\nu_i(N)$ are independent. Following the same procedure as before for canonical

ensemble, we get the expression for the distribution of microstates that maximizes Ω in the form,

$$\ln\left(\frac{M}{\nu_i^*(N)}\right) - \alpha - \beta E_i - \gamma N = 0, \tag{11.8}$$

for each i, or equivalently,

$$\nu_i^*(N) = Me^{-\alpha-\beta E_i-\gamma N}. \tag{11.9}$$

Eq. (11.2) can be used to eliminate M here, namely,

$$M = \sum_i \sum_N \nu_i^*(N) = M \sum_i \sum_N e^{-\alpha-\beta E_i-\gamma N}, \tag{11.10}$$

to yield,

$$e^{-\alpha} = 1/\sum_i \sum_N e^{-\beta E_i-\gamma N}. \tag{11.11}$$

In summary, the probability $p_i(N)$ that the system described by grand canonical ensemble will have N particles and will be in the energetic state $E_i(N,V)$ is,

$$p_i(N) = \lim_{M\to\infty} \frac{\nu_i^*(N)}{M} = \frac{1}{\Xi} e^{-\beta E_i-\gamma N}, \tag{11.12}$$

where,

$$\Xi = \sum_i \sum_N e^{-\beta E_i(N,V)-\gamma N}, \tag{11.13}$$

is the normalization constant that is called as the grand thermodynamic partition function. Still the result is not entirely satisfactory because the Lagrange multiplier γ is not determined. To do that, let us explore thermodynamic connection.

11.2 Thermodynamic properties in grand canonical ensemble

The internal energy and pressure are obtained in close similarity to our previous derivation,

$$U = \sum_i \sum_N E_i(N,V) p_i(N), \tag{11.14}$$

and

$$P = -\sum_i \sum_N p_i(N) \left(\frac{\partial E_i}{\partial V}\right)\Big|_N. \tag{11.15}$$

In addition, the average number of particles that corresponds to the external chemical potential is,

$$\overline{N} = \sum_i \sum_N N p_i(N). \tag{11.16}$$

Let us calculate the differential of U using the explicit expression for $p_i(N)$ from Eq. (11.12),

$$\begin{aligned} dU &= \sum_i \sum_N p_i(N) dE_i(N,V) + \sum_i \sum_N E_i(N,V) dp_i(N) \\ &= \sum_i \sum_N E_i(N,V) dp_i(N) + \sum_i \sum_N p_i(N) \left(\frac{\partial E_i(N,V)}{\partial V}\right) dV \\ &= \sum_i \sum_N E_i(N,V) dp_i(N) - PdV. \end{aligned} \tag{11.17}$$

On the other hand explore the form,

$$-k_B T d(\sum_i \sum_N p_i(N) \ln p_i(N)) = -k_B T \sum_i \sum_N [-\beta E_i(N,V) - \gamma N - \ln \Xi] dp_i(N)$$
$$= \sum_i \sum_N E_i(N,V) dp_i(N) + k_B T \gamma \sum_i d \sum_N N p_i(N) = \sum_i \sum_N E_i(N,V) dp_i(N) + k_B T \gamma d\overline{N}. \tag{11.18}$$

Use now Eq. (11.18) to put Eq. (11.17), in pretty thermodynamic form,

$$dU = -k_B T d(\sum_i \sum_N p_i(N) \ln p_i(N)) - PdV - k_B T \gamma d\overline{N}, \tag{11.19}$$

and compare this expression with thermodynamic formula,

$$dU = TdS - PdV + \mu d\overline{N}. \tag{11.20}$$

We observe that,

$$\gamma = -\mu / k_B T, \quad \beta = 1/k_B T. \tag{11.21}$$

Then the partition function of the grand canonical ensemble takes its final form,

$$\Xi = \sum_i \sum_N e^{-E_i/k_B T + \mu N/k_B T} = \sum_N e^{\mu N/k_B T} Q_N(V,T). \tag{11.22}$$

It can be seen that the grand canonical ensemble can be considered as a set of canonical ensembles with different number of particles. As it follows from our considerations above, the entropy in the grand canonical ensemble is,

$$S = -k_B \sum_i \sum_N p_i(N) \ln p_i(N). \tag{11.23}$$

It is necessary to put emphasis on the properties intrinsic for the grand canonical ensemble. Namely, we can find the probability $p(N)$ that a system with volume V at temperature T has the number of particles N regardless its energy states,

$$p(N) = \sum_i p_i(N) = \frac{Q_N}{\Xi} e^{\mu N/k_B T}. \tag{11.24}$$

Then, the average number of particles in the system with given chemical potential μ and V, T can be written as,

$$\overline{N} = \sum_i \sum_N N p_i(N) = \sum_N N \sum_i p_i(N) = \sum_N N p(N) = \sum_N N \frac{Q_N}{\Xi} e^{\mu N/k_B T}. \tag{11.25}$$

Taking into account Eqs. (11.22) and (11.24), the average number of particles is related to the external variable of the chemical potential as follows,

$$\overline{N} = k_B T \left(\frac{\partial \ln \Xi}{\partial \mu} \right) \Big|_{V,T}. \tag{11.26}$$

Similarly, the equation of state given by Eq. (11.15) is,

$$P = k_B T \left(\frac{\partial \ln \Xi}{\partial V} \right) \Big|_{\mu,T}. \tag{11.27}$$

Finally, we would like to express entropy of the system in terms of the partition function. To do that, let us use the definition of entropy given in Eq. (11.23) and just substitute the result for the probability into $\ln p_i(N)$. Then we obtain,

$$S = \frac{U}{T} - \frac{\mu \overline{N}}{T} + k_B \ln \Xi. \tag{11.28}$$

On the other hand, the formula,

$$S = \left(\frac{k_B \partial (T ln \Xi)}{\partial T}\right)\Big|_{\mu,V}, \tag{11.29}$$

yields the same result.

Finally, it is of interest to make comparison of the expression for entropy in Eq. (11.28) with thermodynamic result in Eq. (11.20), $dU = TdS - PdV + \mu d\overline{N}$, written as,

$$dS = \frac{1}{T}dU + \frac{1}{T}PdV - \frac{1}{T}\mu d\overline{N}. \tag{11.30}$$

We assume that the system is macroscopic, or in other words the contribution of the surface boundary to thermodynamic properties is negligible, comparing to the contribution coming from the body volume of the system. Thus we integrate the expression by volume (taking into account that U is the extensive variable) to get,

$$S = \frac{1}{T}U + \frac{1}{T}PV - \frac{1}{T}\mu\overline{N}. \tag{11.31}$$

It follows then, the equation of state can be written as,

$$PV = k_B T \ln \Xi = -\Omega(\mu, V, T), \tag{11.32}$$

where Ω in thermodynamics is called as the grand thermodynamic potential. It is worth mentioning that this expression is valid for homogeneous macroscopic system, i.e. the surface effects should be vanishing [2].

11.3 Fluctuations of density in grand canonical ensemble

Define and transform the mean square deviation of the number of particles as,

$$\begin{aligned}\sigma_N^2 &= < N^2 > - < N >^2 = \sum_i \sum_N N^2 p_i(N) - < N >^2 \\ &= \frac{1}{\Xi}\sum_N N^2 e^{\beta\mu N} Q_N(V,T) - < N >^2 .\end{aligned} \tag{11.33}$$

We observe that,

$$\begin{aligned}\frac{1}{\Xi}\sum_N N^2 e^{\beta\mu N} Q_N(V,T) &= \frac{1}{\beta^2}\frac{1}{\Xi}\left(\frac{\partial^2 \Xi}{\partial \mu^2}\right) = \frac{1}{\beta^2 \Xi}\frac{\partial}{\partial \mu}\left(\frac{\partial \Xi}{\partial \mu}\right) \\ &= \frac{1}{\beta\Xi}\frac{\partial}{\partial\mu}(\Xi\overline{N}) = \frac{\overline{N}}{\beta\Xi}\left(\frac{\partial \Xi}{\partial \mu}\right) + \frac{1}{\beta}\left(\frac{\partial \overline{N}}{\partial \mu}\right),\end{aligned} \tag{11.34}$$

where the derivatives are taken at constant V and T. Taking into account Eq. (11.26), we arrive at the result,

$$\sigma_N^2 = k_B T \left(\frac{\partial \overline{N}}{\partial \mu}\right). \tag{11.35}$$

Let us derive auxiliary relation by using Eq. (11.27) and chain rule,

$$\begin{aligned}P &= k_B T \left(\frac{\partial ln \Xi}{\partial V}\right)_{\mu,T} = k_B T \left(\frac{\partial ln \Xi}{\partial \mu}\right)_{V,T}\left(\frac{\partial \mu}{\partial <N>}\right)_{V,T}\left(\frac{d < N >}{dV}\right) \\ &= <N> \frac{<N>}{V}\left(\frac{\partial \mu}{\partial <N>}\right)_{V,T} = \frac{<N>^2}{V}\left(\frac{\partial \mu}{\partial <N>}\right)_{V,T},\end{aligned} \tag{11.36}$$

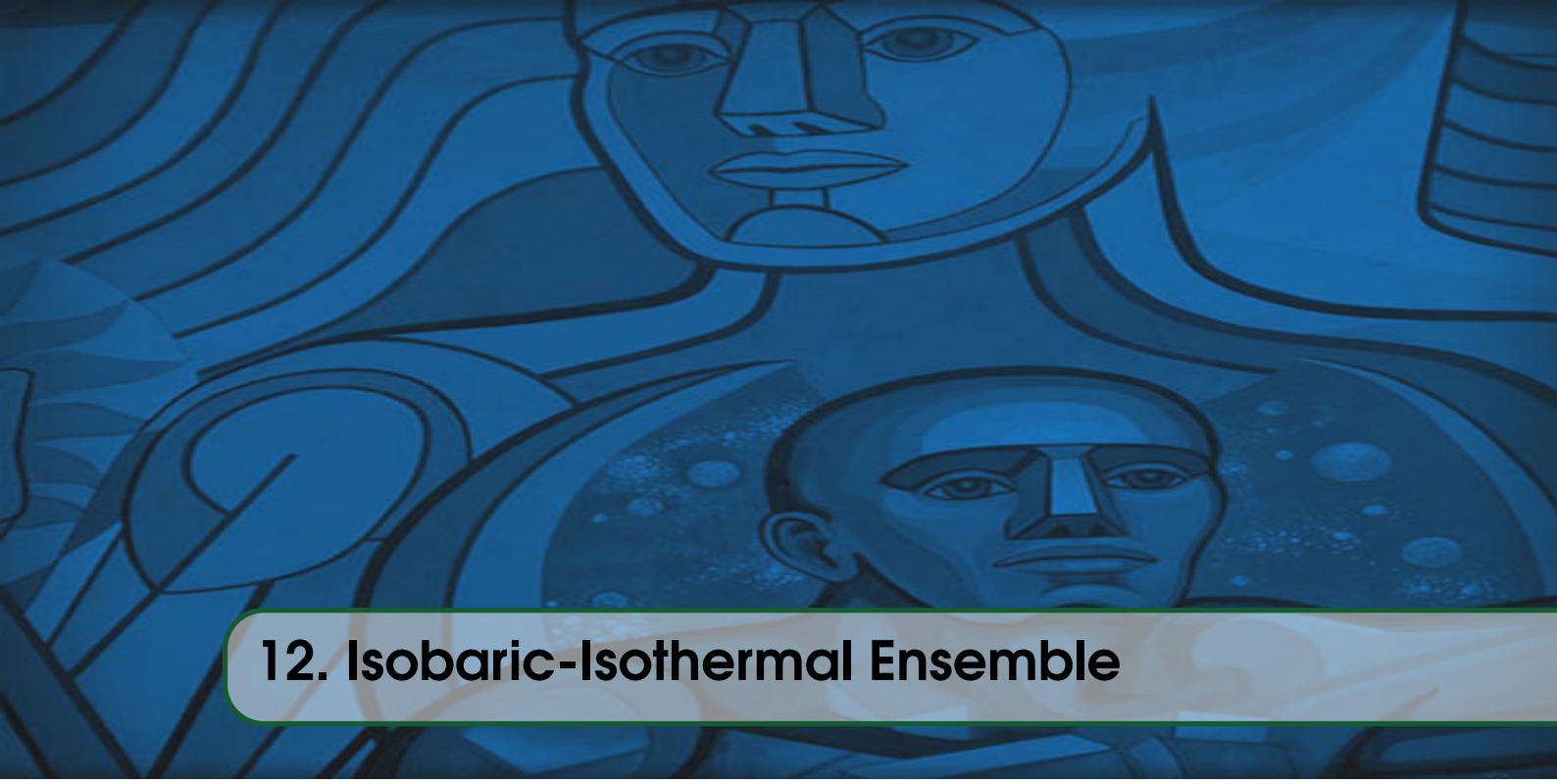
where $<N>$ is defined by Eq. (11.26), and $<N>V$ are extensive variables. So, $d<N>/dV = <N>/V$. The isothermal compressibility, κ_T, follows from Eq. (11.36)

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = \frac{V}{<N>^2}\left(1/\left(\frac{\partial \mu}{\partial <N>}\right)_{V,T}\right). \tag{11.37}$$

In consequence, the fluctuations of density for a system (Eqs. (11.33) and (11.35)) described by the grand canonical ensemble are related to the isothermal compressibility as follows,

$$\sigma_N^2 = \frac{\overline{N}^2}{V} k_B T \kappa_T. \tag{11.38}$$

It is known that at vapor-liquid critical point $(\partial P/\partial V)_T = 0$, or vice versa, $(\partial V/\partial P)_T \to \infty$ while approaching the critical point. As a result we see that the fluctuations of density diverge at the critical point.



12. Isobaric-Isothermal Ensemble

12.1 The Isobaric-isothermal ensemble

Now we would like to consider an ensemble of systems with fixed N, T and P conditions. Under such circumstances the volume of the system, V is the result of assumed value of pressure P. This is the so-called isobaric-isothermal ensemble.

As we know already, the energetic state is determined by the number of particles and volume of the system, e.g. for the i-th state we have $E_i = E_i(N, V)$. For purposes of the derivation, as in the book of T. Davis [2] and in other sources, we would like to discretize the volume rather than use it as a continuous variable. So, we imagine that the volume of the system is $V_j = j \times v_0$, where v_0 is a small volume element, or say unit of measure of the volume. Thus, we can write $E_i = E_i(N, V_j)$ and speak about the energy state i of the system with volume V_j and with number of particles N.

Consider a system at N, P, T conditions and make its M copies, in close similarity to previous chapters. The values for volume variable for each copy (conjugated to the pressure) are unknown, we can suspect that volume will fluctuate. All the NPT systems together, will form a single super-system, next isolate this super-system. The isolated super-system consists of M copies,

$$M = \sum_i \sum_{V_j} \nu_i(V_j), \tag{12.1}$$

where $\nu_i(V_j)$ is the number of systems at energetic state i with the volume V_j. Because of isolation, the super-system is characterized by the total energy,

$$E_{tot} = \sum_i \sum_{V_j} \nu_i(V_j) E_i(N, V_j). \tag{12.2}$$

One more constraint is that the total volume of the isolated super-system is,

$$V = \sum_i \sum_{V_j} \nu_i(V_j) V_j. \tag{12.3}$$

Next, we can construct the microcanonical ensemble of super-systems, each with N, V and E_tot parameters. We have everything now to apply all the machinery as in previous two

chapters at arrive at the probability that the system is in the energetic state $E_i(N, V_j)$ with the volume $E_i(N, V_j)$,

$$p_i(V_j) = \frac{1}{\Delta} e^{-\beta E_i(N,V_j) - \gamma V_j}, \tag{12.4}$$

where β and γ are Lagrange multipliers and Δ is the partition function of the isobaric-isothermal ensemble,

$$\Delta = \sum_i \sum_{V_j} e^{[-\beta E_i(N,V_j) - \gamma V_j]}. \tag{12.5}$$

One can apply similar tricks as for canonical and grand canonical ensemble, by comparison with known result for energy and by comparison with thermodynamic relations, to show that,

$$\beta = 1/k_B T; \quad \gamma = P/k_B T. \tag{12.6}$$

Thus, the partition function is,

$$\Delta = \sum_i \sum_{V_j} e^{-(E_i(N,V_j) - PV_j)/k_B T} = \sum_{V_j} Q_N(V_j) e^{-PV_j/k_B T}, \tag{12.7}$$

where Q_N refers to the partition function of canonical ensemble.

The characteristic thermodynamic potential for the NPT ensemble is the Gibbs free energy, G, in which,

$$G = -k_B T \ln \Delta. \tag{12.8}$$

Although Eq. (12.7) is formally correct, Guggenheim (who introduced the NPT ensemble first) never specified the set of values of V over which the sum was to proceed. In the thermodynamic limit, this omission is inconsequential since almost any reasonable set of volumes is satisfactory. The lack of specification of the volumes over which the sum extends is one of the unsatisfactory features of the NPT ensemble. The success of the NPT ensemble, despite the failure to specify the set of volumes, is closely connected to the properties of all partition functions in the thermodynamic limit. As the system becomes macroscopic in size, Δ can be represented by its maximum term, so in essence one is dealing only with the canonical ensemble. Therefore, the role of the sum in Eq. (12.7) is reduced primarily to facilitate certain mathematical manipulations needed in the derivation of thermodynamic properties.

Since the volumes of most systems are regarded as continuous variables, some authors have attempted to remove the conceptual difficulty associated with the sum over an unspecified set of discrete volumes by expressing Δ as,

$$\Delta = \frac{1}{v_0} \int_V Q_N(V, T) e^{-PV/k_B T} dV. \tag{12.9}$$

The replacement of the sum in Eq. (12.7) by an integral enables the inclusion of all volumes, but at the expense of generating a partition function that has the dimensions of volume. Consequently, this partition function must be rendered dimensionless through division by some constant "quantum" of volume denoted by v_0 Eq. (12.9). Note that the original, and formally correct, partition function in Eq. (12.7) is dimensionless. Therefore, Eq. (12.9), which approximates Eq. (12.7) via the replacement of the summation by an integral, must also be dimensionless. Partition functions, which by definition are pure numbers i.e., sums over states, are related to thermodynamic potentials via a logarithm and so should not have dimensions. The constant v_0, however, cancels out when determining the ensemble average of a given variable and so need not be specified.

The Eq. (12.9) permits to state that the partition function of the NPT ensemble just is the Laplace transform of the partition function of the corresponding canonical ensemble. Thermodynamic properties of the NPT ensemble follow from common thermodynamic formula, it is necessary just to substitute G by the partition function given in Eq. (12.8). One example refers to the calculation of the average volume of the system described by the NPT ensemble,

$$\overline{V} = -\left(\frac{\partial ln\Delta}{\partial \beta P}\right)\Big|_{\beta,N}. \tag{12.10}$$

Exercise 12.1 Show that $< H^2 > - < H >^2 = k_B T^2 C_P$ in an NPT ensemble. Here $H = U + PV$ is enthalpy and $C_P = (\partial < H > /\partial T)|_{N,P}$ is the heat capacity at constant pressure.

Answer: In the isothermal-isobaric ensemble the partition function is given by,

$$\Delta = \sum_V \sum_j e^{-\beta E_j - \beta PV}, \tag{12.11}$$

where $E_j's$ represents the energies of the system at a given volume V. Multiplying both sides of Eq. (12.11) by $H = U + PV$, we obtain,

$$\overline{H}\Delta = \sum_{V,j} (U_j + PV) e^{-\beta E_j - \beta PV}. \tag{12.12}$$

This allows us to take the temperature derivative at constant N and P, as follows,

$$\left(\frac{\partial \overline{H}}{\partial T}\right)_{N,P} \Delta + \overline{H}\left(\frac{1}{k_B T^2}\sum_{V,j}(U_j + PV)e^{-\beta E_j - \beta PV}\right) = \frac{1}{k_B T^2}\sum_{V,j}(U_j + PV)^2 e^{-\beta E_j - \beta PV}, \tag{12.13}$$

next, dividing through by the partition function Δ, we get,

$$\left(\frac{\partial \overline{H}}{\partial T}\right)_{N,P} + \frac{1}{k_B T^2}\overline{H}\underbrace{\frac{\left(\sum_{V,j}(U_j + PV)e^{-\beta E_j - \beta PV}\right)}{\Delta}}_{\overline{H}} = \frac{1}{k_B T^2}\underbrace{\frac{\sum_{V,j}(U_j + PV)^2 e^{-\beta E_j - \beta PV}}{\Delta}}_{\overline{H^2}}, \tag{12.14}$$

this leads to the following relationship,

$$\overline{H^2} - \overline{H}^2 = k_B T^2 \left(\frac{\partial \overline{H}}{\partial T}\right)_{N,P}, \tag{12.15}$$

since $\left(\frac{\partial \overline{H}}{\partial T}\right)_{N,P} = C_p$, we have,

$$\overline{H^2} - \overline{H}^2 = k_B T^2 C_p. \tag{12.16}$$

■

12.2 Classical mechanical limit

Frequently, some or all of the degrees of freedom of a system of particles can be considered in the framework of classical mechanics. This happens if the spacing between neighboring

energy levels is very small comparing to k_BT or in other words, the energy spectrum is continuum. Therefore, we would like to study how to make conversion of the concepts considered in previous chapters to classical mechanics.

To begin with, let us introduce the probability, $P_{E,\delta E}$ that an isothermal system has energy in the interval between E and $E+\delta E$, with δE infinitesimal. The probability is,

$$P_{E,\delta E} = \frac{1}{Q_N}\sum_{i_{low}}^{i_{up}} e^{-\beta E_i}, \tag{12.17}$$

where i_{low} and i_{up} are the numbers delimiting the energy interval in question, $E < E_i < E+\delta E$. As it was mentioned, the interval is small, so we can write the expression above in terms of the number of states in the interval, $W(E;\delta E)$,

$$P_{E,\delta E} = \frac{1}{Q_N} W(E;\delta E) e^{-\beta E}. \tag{12.18}$$

Consider now the same situation from the point of view of classical mechanics. The state of the system is determined by specifying the coordinates and momenta of all the particles, i.e. $\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N$ and $\mathbf{p}_1, \mathbf{p}_2, .., \mathbf{p}_N$. The energy, or equivalently the classical Hamiltonian of the system, H, is a continuous function of the coordinates and momenta. At the beginning of the course, we have noted already that for each pair of the variables q and p the number of *quantum* states in the $(dqdp)$ element of the *classical* phase space is,

$$\frac{1}{h} dqdp. \tag{12.19}$$

Therefore, the number of quantum states in the interval H and $H+\delta H$ for a classical system of N particles is proportional to,

$$\frac{1}{h^{3N}} d^3r_1 ... d^3r_N d^3p_1 .. d^3p_N. \tag{12.20}$$

However, there is one more difference in counting states in quantum and classical mechanics. In quantum mechanics, the states differing only by permutation of mechanically identical particles are not distinct states. Therefore, to avoid over-counting of states in classical mechanics, to conform with quantum mechanics, one should modify the expression above dividing it by $N!$. Hence, the so-called correspondence principle can be formulated as follows,

$$\frac{1}{Q_N} W(E;\delta E) e^{-\beta E} \to \frac{1}{Q_N} \frac{e^{-\beta H}}{N! h^{3N}} d^3r_1 .. d^3p_N \equiv P_N(\mathbf{r}_1, .., \mathbf{p}_N) d^3r_1 ... d^3r_N d^3p_1 .. d^3p_N, \tag{12.21}$$

where $P_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{p}_N) d^3r_1 ... d^3r_N d^3p_1 .. d^3p_N$ is the probability to find N particles in the interval of coordinates and momenta $\mathbf{r}_i, \mathbf{r}_i + d\mathbf{r}_i$ and $\mathbf{p}_i, \mathbf{p}_i + d\mathbf{p}_i$ and $P_N(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{p}_N)$ is the probability density. If the r.h.s. of equation above is valid, we say that the system obeys classical mechanics and Boltzmann statistics.

Normalization condition for the probability density,

$$\int ... \int d^3r_1 ... d^3r_N d^3p_1 .. d^3p_N P_N(\mathbf{r}_1, .., \mathbf{p}_N) = 1, \tag{12.22}$$

immediately leads to the expression for the partition function of the classical system,

$$Q_N(V,T) = \frac{1}{N! h^{3N}} \int ... \int d^3r_1 ... d^3r_N d^3p_1 .. d^3p_N e^{-\beta H}. \tag{12.23}$$

In general, the Hamiltonian of the classical system of N particles is the sum of kinetic and potential energy,

$$H = \sum_i \frac{p_i^2}{2m_i} + U_N(\mathbf{r}_1, .., \mathbf{r}_N). \tag{12.24}$$

Then the partition function can be written as follows,

$$Q_N(V,T) = \frac{1}{N!h^{3N}} \int ... \int d^3p_1 ... d^3p_N e^{-\beta \sum_i \frac{p_i^2}{2m_i}} \int ... \int d^3r_1 ... d^3r_N e^{-\beta U_N(\mathbf{r}_1,..,\mathbf{r}_N)}$$
$$= \frac{1}{N!} \left(\frac{2\pi m_i k_B T}{h^2} \right)^{3N/2} \int ... \int d^3r_1 ... d^3r_N e^{-\beta U_N(\mathbf{r}_1,..,\mathbf{r}_N)}. \tag{12.25}$$

If the system is ideal, $U_N(\mathbf{r}_1,..,\mathbf{r}_N) = 0$, we obtain the partition function of a classical ideal gas described by canonical ensemble,

$$Q_N^{id} = V^N \left(\frac{2\pi m_i k_B T}{h^2} \right)^{3N/2}. \tag{12.26}$$

Then, general expression for a non-ideal fluid system takes the form,

$$Q_N(V,T) = Q_N^{id} Z_N(V,T), \tag{12.27}$$

where,

$$Z_N(V,T) = \frac{1}{V^N} \int ... \int d^3r_1 ... d^3r_N e^{-\beta U_N(\mathbf{r}_1,..,\mathbf{r}_N)}, \tag{12.28}$$

is called as the configurational partition function because it comes from configurations of particles determined by the inter-particle interactions.

There exists one additional peculiarity while counting states within classical statistics. If our system is a classical solid, the particles are distinguishable because they can be identified by the coordinates of the lattice sites they occupy. In consequence, it is not necessary to introduce the factor $1/N!$ into the expression for the corresponding partition function. We will illustrate slightly below. Prior doing that, let us just comment the following issue. In many circumstances some of the degrees of freedom of every molecule, for example the ones describing translational and rotational motion, behave classically, whereas other degrees of freedom, like vibrational motion and electronic degrees of freedom, require quantum mechanical description. It is then assumed that the Hamiltonian consists of two independent parts. For example $H^{(1)}$ is chosen to represent the classical mechanical degrees of freedom and is assumed to be independent on $H^{(2)}$ that describes the quantum mechanical degrees of freedom. Then the partition function should be written and calculated as the product of two contributions,

$$Q_N(V,T) = \left(\sum_i e^{-\beta E_i^{(2)}} \right) \times \frac{1}{N!h^{3N}} \int ... \int d^3q_1 ... d^3q_N d^3p_1 ... d^3p_N e^{-\beta H^{(1)}}, \tag{12.29}$$

where N should correspond to the number of classical degrees of freedom.

Now, let us illustrate the correspondence principle for the correct counting of states upon considering classical and quantum partition functions for a solid. Specifically what we like to do is to find the expression for the coefficient C_N that enters the relation,

$$\sum_i e^{-\beta E_i} \to C_N \int ... \int d^3q_1 ... d^3q_N d^3p_1 ... d^3p_N e^{-\beta H}, \tag{12.30}$$

by considering the so-called Einstein's model for a solid.

This model for a solid involves a set of N independent, identical but distinguishable three-dimensional harmonic oscillators such that the Hamiltonian of the problem is,

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i=1}^{N} \frac{\lambda}{2} (\mathbf{r}_i - \mathbf{R}_i)^2, \tag{12.31}$$

where $\mathbf{R}_i$ gives a set coordinates of lattice sites.

The quantum mechanical partition function for the problem in question can be obtained because the energy levels of a single quantum harmonic oscillator are available from Eq. (8.55). The partition function is just the product for the sum over states of independent oscillators, say a quantum ideal gas without any relevance to the coordinate space, and explicitly is given in the form,

$$\begin{aligned}\sum_i e^{-\beta E_i} &= \sum_{\nu_1=0}^{\infty} ... \sum_{\nu_{3N}=0}^{\infty} \exp\left[-\frac{\hbar\omega}{k_B T}(\nu_1 + 1/2)\right] ... \exp\left[-\frac{\hbar\omega}{k_B T}(\nu_{3N} + 1/2)\right] \\ &= \frac{\exp\left[-\frac{3N\hbar\omega}{2k_B T}\right]}{\left(1 - \exp[-\frac{\hbar\omega}{k_B T}]\right)^{3N}}. \end{aligned} \tag{12.32}$$

The frequency of each oscillator is $\omega = (\lambda/m)^{1/2}$. On the other hand, a part of the partition function (without integration over coordinates) coming from the kinetic energy of an ideal gas of classical oscillators is,

$$\int ... \int d^3p_1 ... d^3p_N e^{-\beta H} = \left[(2\pi k_B T)^2 \left(\frac{m}{\lambda}\right)\right]^{3N/2}. \tag{12.33}$$

Thus,

$$\left(\frac{\exp[-\frac{\hbar\omega}{2k_B T}]}{(1 - \exp[-\frac{\hbar\omega}{k_B T}])}\right)^{3N} \to C_N \left[(2\pi k_B T)\left(\frac{m}{\lambda}\right)^{1/2}\right]^{3N}. \tag{12.34}$$

The l.h.s. of this expression should tend to the classical result in the limit of high temperature or in other words when $\hbar\omega/k_B T$ becomes small (this issue has been discussed in the chapter concerned with the applicability of either quantum or classical treatment of the system in terms of degeneracy temperature). Thus, we perform Taylor expansion of the terms on the l.h.s. to arrive at,

$$\left[\frac{k_B T}{\hbar\omega}(1 - \frac{1}{24}(\frac{\hbar\omega}{k_B T})^2 + ...)\right]^{3N} \to C_N \left[(2\pi k_B T)\left(\frac{m}{\lambda}\right)^{1/2}\right]^{3N}. \tag{12.35}$$

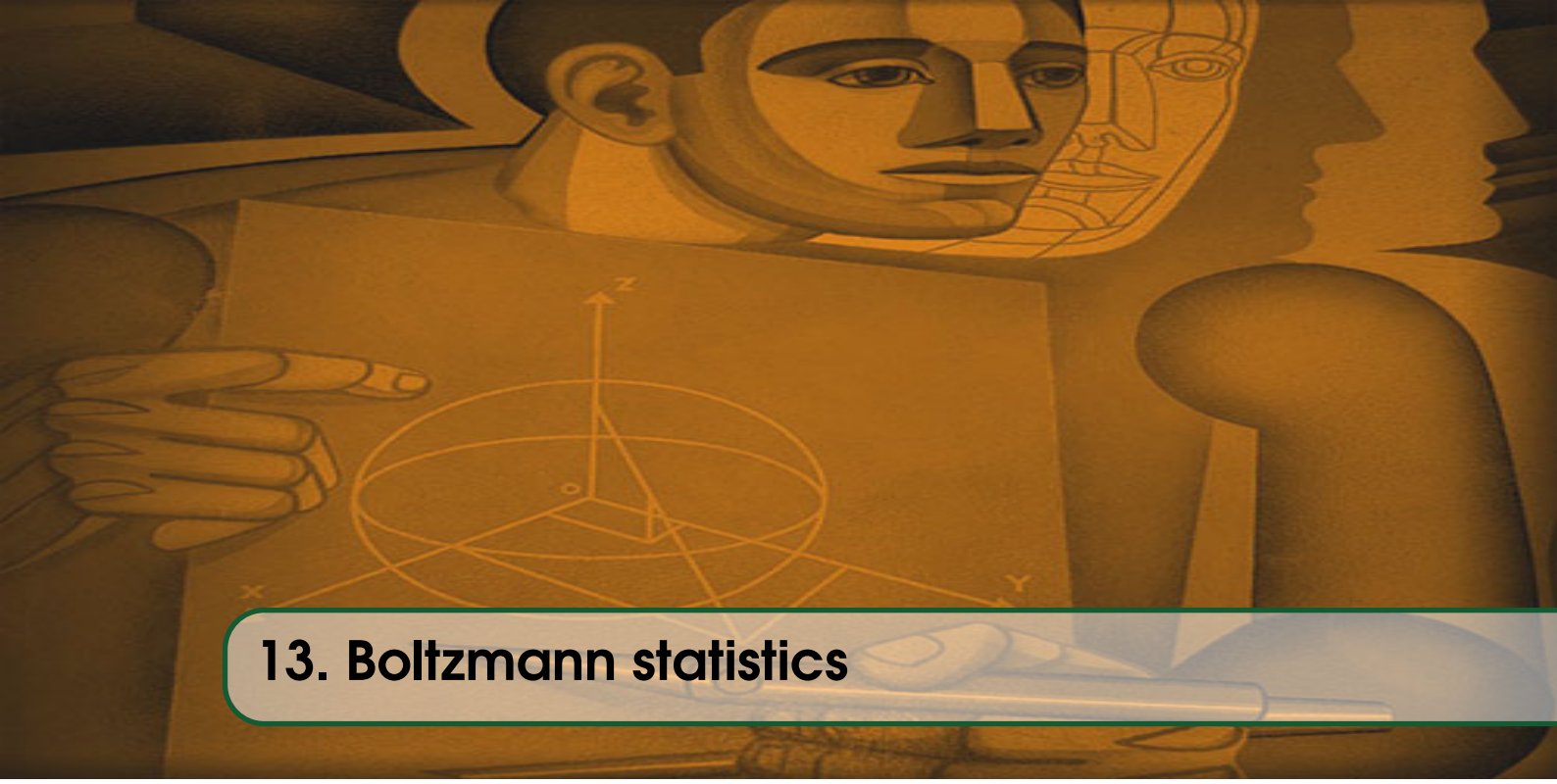
In consequence,

$$C_N = \frac{1}{h^{3N}}, \tag{12.36}$$

in agreement with our discussion of the correspondence principle in this subsection above.

Exercise 12.2 Show that 1/24 is the correct multiplier in the expansion. ■

Exercise 12.3 Write down equations describing thermodynamic properties of a system described by microcanonical ensemble. ■



13. Boltzmann statistics

13.1 The Boltzmann statistics

In the previous chapters, cf. chapter 10, we have given the expressions for the probability distribution function and for the partition function of the system described by canonical ensemble (NVT) by using Boltzmann statistics. A more detailed description is provided below with examination of the properties of ideal gas.

Let us start from generalities. Namely, consider an ideal gas of identical particles. Let the system of N particles is in a specific quantum state, $\{j\}$, described by a set of numbers $(j_1, j_2, ..., j_N)$. Then the energy of the system in such state is,

$$E_{\{j\}} = \varepsilon_{1j_1} + ... + \varepsilon_{Nj_N} = \sum_{i=1}^{N} \varepsilon_{ij_i}, \tag{13.1}$$

where ε_{ij_i} is the energy if the i-th particle in one-particle quantum state j_i. According to the definition of the partition function, we have,

$$Q_N(V,T) = \sum_{\{j\}} \exp\left\{-\beta \sum_{i=1}^{N} \varepsilon_{ij_i}\right\}, \tag{13.2}$$

where the summation over Boltzmann multipliers (the first sum in the r.h.s.) is performed over all **distinct** quantum states. How this summation applies? The particles of an ideal gas are non-interacting, thus one might expect that the summation over quantum states of the entire system can be performed independently, i.e. over one-particle states,

$$Q_N(V,T) = \left(\sum_{j_1} \exp\{-\beta\varepsilon_{1j_1}\}\right) \times ... \times \left(\sum_{j_N} \exp\{-\beta\varepsilon_{Nj_N}\}\right) = \prod_{i=1}^{N} q_i, \tag{13.3}$$

where q_i is the partition function of a single particle i. Two important comments at this point are the following:

1) If the particles were not identical (i.e. possessing certain different labels) they would be distinguishable and consequently Eq. (13.3) for Q_N would be correct.

2) However, if the particles are identical, they are indistinguishable according to quantum mechanics. As a result, the quantum states that differ **only** by the interchange

of identical particles **among the same set** of one-particle energy states are identical or in other words are not distinct.

Consequently, the product in the form q^N overestimates the partition function Q_N, because non-distinct quantum states are included in q^N.

■ **Example 13.1** Assume that we have a system of two identical particles 1 and 2 in which there are only two one-particle states a and b. Introduce notation, $\varepsilon_{i,a}$ and $\varepsilon_{i,b}$ with $i = 1,2$ for the energy states of particles. According to Eq. (13.2), the partition function is,

$$Q_2 = \exp\{-\beta(\varepsilon_{1,a} + \varepsilon_{2,a})\} + \exp\{-\beta(\varepsilon_{1,b} + \varepsilon_{2,b})\} + \exp\{-\beta(\varepsilon_{1,a} + \varepsilon_{2,b})\}. \tag{13.4}$$

The term with $\varepsilon_{1,b} + \varepsilon_{2,a}$ does not appear because it yields the state identical to $\varepsilon_{1,a} + \varepsilon_{2,b}$. However, the product,

$$\begin{aligned} q^2 &= \exp\{-\beta(\varepsilon_{1,a} + \varepsilon_{2,a})\} + \exp\{-\beta(\varepsilon_{1,b} + \varepsilon_{2,b})\} \\ &+ \exp\{-\beta(\varepsilon_{1,a} + \varepsilon_{2,b})\} + \exp\{-\beta(\varepsilon_{1,b} + \varepsilon_{2,a})\} > Q_2. \end{aligned} \tag{13.5}$$

In general, evaluation of Q_N will be a difficult procedure of assessing each state for indistinguishability. As we discussed in the previous chapter 12, Q_N can be adequately approximated by,

$$Q_N = \frac{1}{N!} q^N, \tag{13.6}$$

for macroscopic systems such as gases or liquids at not very low temperature. Actually, $N!$ corresponds to the number of ways N particles can be permuted among all different one-particle states. We would like to note that if Eq. (13.6) is valid, we say that the system obeys Boltzmann or classical statistics. ■

Comment: In the following applications we would require generalization of Eq. (13.6) for mixtures. Denote a as the species index and the number of components as M,

$$\sum_{a=1}^{M} N_a = N. \tag{13.7}$$

The partition function of the system then is,

$$Q_N = \prod_{a=1}^{M} \frac{1}{N_a!} q_a^{N_a}, \tag{13.8}$$

where q_a is the one-particle partition function belonging to species a.

Return now to a one-component gas, but assume that molecules have internal structure. The translational motion of the center of mass of the molecules is always independent of the internal degrees of freedom. We have seen one example already in chapter 9 while studying the form of kinetic energy of the diatomic molecule. With such decoupling, we can formally write down the partition function of a molecule as $q = q_{tr} q_{int}$, where the second multiplier is the partition function corresponding to the internal degrees of freedom (rotational, vibrational, nuclear and electronic). The partition function of internal degrees of freedom depends on temperature but does not depend on volume. The translational motion is well described by classical mechanics. Then, integrating kinetic energy of the center of mass, according to the definition of the partition function it is easy to obtain the one-particle result as,

$$q^{(tr)} = \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} V, \tag{13.9}$$

where m is the molecular mass. Recall that $A = -k_B T \ln Q_N$, cf. Eq. (10.65). Then, applying Eq. (13.6) and the result above, one arrives at the equation of state of ideal gas,

$$P = -\left(\frac{\partial A}{\partial V} \right) = N \left(\frac{\partial \ln q_{tr}}{\partial V} \right) = \frac{N k_B T}{V}. \tag{13.10}$$

In addition, the chemical potential can be found from,

$$\mu = \left(\frac{\partial A}{\partial N}\right) = k_B T \ln\left(\frac{N}{V}\right) + \mu^*(T), \tag{13.11}$$

where

$$\mu^*(T) = -k_B T \ln\left[q_{int}(T)\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2}\right]. \tag{13.12}$$

It is of interest to explore conditions under which Boltzmann statistics holds. In other words, we would like to extract physical information from the presence of the multiplier $1/N!$ in Eq. (13.6). As we have mentioned slightly above, this multiplier provides correct counting of states in the situation in which individual particles are in different one-particle energetic states. The probability that a particle is in the state with energy ε is proportional to $\exp(-\beta\varepsilon)$. If energy is big, the probability is quite low. On the other hand, if ε by value is below or near $k_B T$ the states can be termed as "easily accessible" because the probability is higher than in former situation. Therefore, if $\Phi(\varepsilon)$ denotes the number of quantum states with energy less and equal to ε, it is sufficient to require

$$N << \Phi(\varepsilon = k_B T). \tag{13.13}$$

We will have then, that the occupation of a single-particle level is below and permutations of identical particles are properly accounted for, because all N particles are almost surely in different quantum states. One can obtain a lower bound for $\Phi(\varepsilon)$ by considering only the translational motion. The presence of internal energy states would only strengthen the inequality above. The translational energy levels of a particle in a cubic box are determined by,

$$\varepsilon = \frac{h^2}{8mV^{2/3}}(n_x^2 + n_y^2 + n_z^2), \tag{13.14}$$

where n_x, n_y and n_z are positive integers. Thus, we can define a sphere in three dimensions with the radius R_ε,

$$R_\varepsilon^2 = n_x^2 + n_y^2 + n_z^2 = \frac{8mV^{2/3}}{h^2}\varepsilon, \tag{13.15}$$

and calculate the number of quantum states with energy between 0 and ε,

$$\Phi(\varepsilon) = \frac{1}{8}\left(\frac{4}{3}\pi R_\varepsilon^3\right) = \frac{\pi}{6}\left(\frac{8m\varepsilon}{h^2}\right)^{3/2} V. \tag{13.16}$$

Hence, the condition of validity of Boltzmann statistics given by Eq. (13.13) can be expressed as follows,

$$\frac{N}{V}\lambda^3 << 1, \tag{13.17}$$

where λ is the de Broglie thermal wavelength, $\lambda = h/\sqrt{2\pi m k_B T}$. Note, that the factor of the order of unity has been omitted upon transformation above. Our result means that Boltzmann statistics is valid, if the de Broglie thermal wavelength is much smaller than the average distance between particles or in other words matter would behave as discrete particles instead of overlapping waves. Also, it is worth mentioning that the condition given by Eq. (13.17) is equivalent to what we have obtained using different arguments in chapter 7, see Eq. (7.9). In addition, it appears that if temperature is very low and the system to study is composed of very light particles, e.g. electrons, one should develop alternatives to Eq. (13.6) of this chapter. These developments lead to Fermi-Dirac or Bose-Einstein statistics which we will consider in following chapters.

13.2 Partition function and thermodynamic properties of monoatomic gases

The energy of a single monoatomic particle of an ideal gas can be written as a sum of different contributions,

$$\varepsilon = \varepsilon^{(n)} + \varepsilon^{(el)} + \varepsilon^{(tr)}, \tag{13.18}$$

that denote the nuclear, electronic and translational energy of the particle. According to quantum mechanical calculations [20, 21], the spacing between nuclear energy ground state and the first excited state is very big. Thus, for fluids under study, it is assumed that the particles remain mostly in their ground state. Then to a good approximation, the electronic and nuclear states can be considered as independent [2]. Thus, the one-particle partition function is,

$$q = q^{(tr)} q^{(n)} q^{(el)}, \tag{13.19}$$

where

$$q^{(tr)} = \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \exp\left\{ -\frac{\beta h^2}{8mV^{2/3}} \left(n_x^2 + n_y^2 + n_z^2 \right) \right\}, \tag{13.20}$$

and

$$q^{(n)} = \sum_{(nuclear)} e^{-\beta \varepsilon_i^{(n)}}; \quad q^{(el)} = \sum_{(electronic)} e^{-\beta \varepsilon_i^{(e)}}. \tag{13.21}$$

Quantum states are often degenerate, i.e. more than one state can have the same energy. Let introduce notation for the degeneracy of the state with energy ε_i as w_i. Then, the nuclear and electronic contributions to the partition function ($q^{(n)}$ and $q^{(el)}$) can be written as,

$$q^{(n)} = \sum_{energy\ levels} w_{ni} e^{-\beta \varepsilon_i^{(n)}} = e^{-\beta \varepsilon_1^{(n)}} \sum_{energy\ levels} w_{ni} e^{-\beta \Delta \varepsilon_i^{(n)}}, \tag{13.22}$$

and,

$$q^{(el)} = \sum_{energy\ levels} w_{ei} e^{-\beta \varepsilon_i^{(e)}} = e^{-\beta \varepsilon_1^{(e)}} \sum_{energy\ levels} w_{ei} e^{-\beta \Delta \varepsilon_i^{(e)}}, \tag{13.23}$$

with the notation $\Delta \varepsilon_i \equiv \varepsilon_i - \varepsilon_1$ both for nuclear and electronic terms.

By convention, the energy of the ground state for nuclei is taken as reference, i.e. $\varepsilon_1^{(n)} = 0$. On the other hand, as it was mentioned above in the discussion of Eq. (13.19), the values of $\varepsilon_i^{(n)} / k_B$ for $i > 1$ are of the order of 10^9 to $10^{10} K$ [2]. Then $e^{-\beta \varepsilon_i^{(n)}} \approx 0$ for $i > 1$ and consequently,

$$q^{(n)} = w_{n1}, \tag{13.24}$$

provides a very good approximation. Concerning the electronic contribution to the partition function, we would like to make the following comments. It has been established that $\Delta \varepsilon_i^{(e)}$ for $i > 1$ often is much greater than unity. Then, a quite good approximation for $q^{(el)}$ is,

$$q^{(el)} = w_{e1} e^{-\beta \varepsilon_1^{(e)}}. \tag{13.25}$$

In the case one needs better accuracy for high temperature of the order $10^3 K$, the first two terms in Eq. (13.23) should be kept,

$$q^{(el)} = \left[w_{e1} + w_{e2} e^{-\beta \Delta \varepsilon_2^{(e)}} \right] e^{-\beta \varepsilon_1^{(e)}}. \tag{13.26}$$

Pertinent parameters, necessary to apply in Eq. (13.26), follow from the analysis of optical spectra and are given in Ref. [22]. These data are frequently used to justify calculations of the partition function of electronic degrees of freedom [2, 7].

Finally, the contribution coming from translational motion is obtained as described in chapter 9, see Eqs. (13.10)- (13.14), to obtain our expression given above in Eq. (13.9). With these results for the partition function of a single monoatomic particle, the partition function for a gas can be written as,

$$Q = \frac{1}{N!}q^N = \frac{1}{N!}(q^{(tr)}q^{(n)}q^{(el)})^N = \frac{(w_{n1}q^{(el)})^N}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2} V^N, \tag{13.27}$$

where $q^{(el)}$ is given either by Eq. (13.25) or Eq. (13.26). Then, the expressions for a set of thermodynamic properties follow straightforwardly. While writing the free energy, $A = -k_B T \ln Q_N$, explicitly we use Stirling approximation $N! = (N/e)^N$ and Eq. (13.26) for the electronic term. Then,

$$\begin{aligned} A = A^{(tr)}A^{(el)}A^{(n)} &= -Nk_BT\ln\left[\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} e\frac{V}{N}\right] \\ -Nk_BT\ln&\left\{\left[w_{e1} + w_{e2}e^{-\beta\Delta\varepsilon_2^{(e)}}\right]e^{-\beta\varepsilon_1^{(e)}}\right\} - Nk_BT\ln w_{n1}. \end{aligned} \tag{13.28}$$

The entropy of the system follows from $S = -(\frac{\partial A}{\partial T})|_{N,V}$, next the internal energy is $U = A + TS$, and finally the heat capacity at constant volume comes out as,

$$C_V = \frac{3}{2}Nk_B + \frac{N}{k_BT^2}(w_{e1}w_{e2})(\Delta\varepsilon_2^{(e)})^2\frac{e^{-\Delta\varepsilon_2^{(e)}/k_BT}}{[w_{e1} + w_{e2}e^{-\Delta\varepsilon_2^{(e)}/k_BT}]^2}. \tag{13.29}$$

The nuclear states contribute to the free energy and entropy due to the degeneracy of the ground state. They do not contribute to the constant volume heat capacity however. The equation of state results from the expression for the free energy in Eq. (13.28), $P = -(\frac{\partial A}{\partial V})|_{N,T}$,

$$P = \frac{Nk_BT}{V}, \tag{13.30}$$

showing that the internal states of the particles of a gas do not contribute into pressure.

It is of interest to explore the case of Boltzmann statistics for a mixture of ideal gases. The partition function for such a case is given by Eq. (13.8). Having in mind that the equation of state is $P = \sum_a N_a k_B T/V$, let us write down the free energy of a mixture of gases as,

$$\begin{aligned} A &= -k_BT\sum_{a=1}^{M}\ln\left(\frac{q_a^{N_a}}{N_a!}\right) = -k_BT\sum_{a=1}^{M}\ln\left(\frac{\tilde{q}_a^{N_a}V^{N_a}e^{N_a}}{N_a^{N_a}}\right) \\ &= -k_BT\sum_{a=1}^{M}\ln\left(\frac{\tilde{q}_a^{N_a}N^{N_a}(k_BT)^{N_a}e^{N_a}}{N_a^{N_a}P^{N_a}}\right) \\ &= -k_BT\sum_{a=1}^{M}\ln\left(\frac{\tilde{q}_a^{N_a}(k_BT)^{N_a}e^{N_a}}{P^{N_a}}\right) + k_BT\sum_{a=1}^{M}N_a\ln\left(\frac{N_a}{N}\right) \\ &= \sum_{a=1}^{M}A_a(N_a,P,T) + k_BT\sum_{a=1}^{M}N_a\ln\left(\frac{N_a}{N}\right). \end{aligned} \tag{13.31}$$

The first term, $A_a(N_a, P, T)$, is the contribution of each species into the free energy of the system whereas the second term,

$$A_{mix} = k_BT\sum_{a=1}^{M}N_a\ln\left(\frac{N_a}{N}\right), \tag{13.32}$$

is the Helmholtz free energy of mixing given in terms of the system's composition, i.e. of species concentrations N_a/N. Then, the entropy and mixing term of entropy are obtained from the derivative of free energy,

$$S = \sum_{a=1}^{M} S_a(N_a, P, T) + S_{mix}, \tag{13.33}$$

where,

$$S_{mix} = -k_B \sum_{a=1}^{M} N_a \ln\left(\frac{N_a}{N}\right). \tag{13.34}$$

The mixing entropy is positive because the logarithms in the r.h.s. are negative. Finally, it is straightforward to see that mixing of species does not contribute into the internal energy.

13.3 Polyatomic gases

Now, we would like to study a more complex system, namely a gas of polyatomic molecules. For simplicity, it is assumed that only the grand state of nuclear and electronic energies contribute to the partition function. Moreover, we restrict attention to the case when the translational center of mass motion as well as the rotational and vibrational motions are independent of one another. Then the partition function of a single molecule has the form,

$$q = q^{(n)} q^{(el)} q^{(tr)} q^{(rot)} q^{(vib)}, \tag{13.35}$$

where the multipliers from left to right refer to nuclear, electronic, translational, rotational and vibrational partition functions, respectively. The first three of them are known from previous analyses, with approximations applied they are,

$$q^{(n)} = w_{n1}; \quad q^{(el)} = w_{e1} e^{-\varepsilon_1^e/k_B T}; \quad q^{(tr)} = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} V. \tag{13.36}$$

Often, the description of rotational and sometimes of vibrational motions can be given adequately using classical mechanics. Explore this version of the procedure. We start from the Hamiltonian for the diatomic molecule considered in chapter 9, $H^{(rot)} = \frac{1}{2} I\omega^2 = \frac{1}{2} I(\dot{\theta}^2 + sin^2\theta\dot{\varphi}^2)$. According to the Hamilton equations of motion,

$$p_\theta = \frac{\partial H^{(rot)}}{\partial \dot{\theta}} = I\dot{\theta}, \quad p_\varphi = \frac{\partial H^{(rot)}}{\partial \dot{\varphi}} = Isin^2\theta\dot{\varphi}, \tag{13.37}$$

such that,

$$H^{(rot)} = \frac{p_\theta^2}{2I} + \frac{p_\varphi^2}{2Isin^2\theta}. \tag{13.38}$$

Then the rotational partition function for a single particle is,

$$q^{(rot)} = \frac{1}{\sigma h^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dp_\theta dp_\varphi \int_0^\pi d\theta \int_0^{2\pi} d\varphi \, \exp\{-H^{(rot)}/k_B T\} = \frac{8\pi^2 I k_B T}{\sigma h^2}. \tag{13.39}$$

The factor $1/h$ results from each pair of variables (θ, p_θ and φ, p_φ). On the other hand, the factor σ comes out from "degeneracy" of the molecule: $\sigma = 1$, if the diatomic molecule is heteronuclear and $\sigma = 2$, if the molecule is homonuclear. Eq. (13.39) gives the classical rotational partition function of any linear polyatomic molecule, just remember that the

effect of vibrations on the momentum of inertia has been neglected, i.e. the rotational and vibrational motion are decoupled.

In the case of a nonlinear polyatomic molecule, the Hamiltonian of rotations is more complicated, then three Euler angles are involved to perform integrations over relevant variables. Details can be found in Ref. [7] and the references therein. The final form of the rotational partition function for such particle is,

$$q^{(rot)} = \frac{(\pi I_A I_B I_C)^{1/2}}{\sigma} \left(\frac{8\pi^2 k_B T}{h^2} \right)^{3/2}, \tag{13.40}$$

where I_A, I_B and I_C are the principal momenta of inertia of the molecule in question. In addition, in such a case σ is the number of identical configurations attained by the rotation of the molecule w.r.t. its center of mass, e.g. $\sigma = 12$ for benzene molecule.

Now, let us proceed to the description of calculation of the vibrational partition function. As we have seen already (see chapter 9) the classical mechanical vibrational energy for a diatomic molecule is,

$$H^{(vib)} = \frac{p^2}{2m_r} + \frac{\lambda x^2}{2} + u_0, \tag{13.41}$$

where m_r is the reduced mass of the molecule, the coordinate x is the deviation of instant position of atoms w.r.t. the center of mass (see Eq. (9.10)). The classical mechanical vibrational partition function then is,

$$q^{(vib)} = \frac{1}{h} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dp \, \exp\{-H^{(vib)}/k_B T\} = \frac{2\pi k_B T}{h} \left(\frac{m_r}{\lambda} \right)^{1/2} e^{-u_0/k_B T}. \tag{13.42}$$

Concerning more complex molecules, it is worth mentioning that the vibrational motions can be quite well approximated as small displacements from the lowest energy configurations. Then, the classical vibrational energy can be well described as a system of coupled harmonic oscillators. Therefore, one needs to apply certain appropriate coordinate transformation (diagonalization of the matrix, see any linear algebra textbook) to convert the Hamiltonian of n_v coupled harmonic oscillators into the form,

$$H^{(vib)} = \sum_{i=1}^{n_v} \left[\frac{p_i^2}{2m_i} + \frac{\lambda_i \xi_i^2}{2} \right] + u_0, \tag{13.43}$$

where the parameters refer to each oscillator. With such transformed Hamiltonian, the vibrational one-particle partition function is,

$$\begin{aligned} q^{(vib)} =& \frac{e^{-u_0/k_B T}}{h^{n_v}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} d\xi_i dp_i \prod_{i=1}^{n_v} \exp\left\{ -\left(\frac{p_i^2}{2m_i} + \frac{\lambda_i \xi_i^2}{2} \right) / k_B T \right\}, \\ =& \left(\frac{2\pi k_B T}{h} \right)^{n_v} \prod_{i=1}^{n_v} \left(\frac{m_i}{\lambda_i} \right)^{1/2} e^{-u_0/k_B T}, \end{aligned} \tag{13.44}$$

where n_v corresponds to the number of vibrational degrees of freedom of the molecule. Recall, that if the molecule consists of n_a atoms, the number of vibrational modes is $n_v = 3n_a - 5$ for a linear molecule and $n_v = 3n_a - 6$ for a nonlinear molecule.

Now, to summarize the developments above, we are able to write down the partition function for a polyatomic molecule and next the Helmholtz free energy of an ideal gas of polyatomic molecules,

$$Q_N = \frac{\left(q^{(t)r} q^{(n)} q^{(e)l} q^{(rot)} q^{(vib)} \right)^N}{N!}, \tag{13.45}$$

and

$$\begin{aligned} A =& A^{(tr)} + A^{(n)} + A^{(el)} + A^{(rot)} + A^{(vib)}, \\ =& - Nk_BT \ln \left[\left(\frac{2\pi M k_B T}{h^2} \right)^{3/2} e \frac{V}{N} \right] - Nk_BT \ln w_{n1} - Nk_BT \ln w_{e1} + N\varepsilon_1^e \\ & - Nk_BT \ln \left(\frac{8\pi^2 I k_B T}{\sigma h^2} \right) - Nk_BT \ln \left[\left(\frac{2\pi k_B T}{h} \right)^{n_v} \prod_{i=1}^{n_v} \left(\frac{m_i}{\lambda_i} \right)^{1/2} \right] + Nu_0. \end{aligned} \quad (13.46)$$

With this expression, the entropy follows from $S = -(\partial A/\partial T)|_{N,V}$ and next, the internal energy is obtained from $U = A + TS$,

$$\begin{aligned} U =& U^{(tr)} + U^{(n)} + U^{(el)} + U^{(rot)} + U^{(vib)}, \\ =& \frac{3}{2} Nk_BT + N\varepsilon_1^e + Nk_BT + n_v Nk_BT + u_0. \end{aligned} \quad (13.47)$$

where we restricted to Eq. (13.39) for the rotational partition function only, generalized form given by Eq. (13.40) can be used as well. The result above in Eq. (13.47) illustrates the so-called equipartition theorem of classical statistical mechanics. Namely, each of translational and rotational degrees of freedom yield the $k_BT/2$ contribution to the internal energy whereas each vibrational mode makes the k_BT contribution.

Finally, it is obvious that the translational contribution to the constant volume heat capacity is $C_V^{(tr)} = 3Nk_B/2$, and other contributions are: $C_V^{(rot)} = Nk_B$ (for diatomic molecules) or $C_V^{(rot)} = 3Nk_B/2$ (for polyatomic molecules). Finally, the term coming from vibrations, $C_V^{(vib)}$, is given by Nk_B multiplied by either $3n_a - 5$ or $3n_a - 6$, dependent on the molecular structure, linear or nonlinear. After analyses of the rotational and vibrational partition functions within classical mechanics, we proceed in the next chapter to the quantum mechanics counterpart of the procedure.

14. Rotational and vibrational partition functions

14.1 Rotational and vibrational partition functions

The quantum mechanical form of the rotational partition function for a single particle is,

$$q^{(rot)} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} e^{\left[-l(l+1)\frac{h^2}{8\pi^2 I k_B T}\right]} = \sum_{l=0}^{\infty} (2l+1) e^{-l(l+1)\Theta_r/T}, \tag{14.1}$$

where the energy eigenvalues have been taken from Eq. (8.35) and the notation $\Theta_r = h^2/8\pi^2 I k_B$ denotes the characteristic rotational energy in temperature units.

The rotational partition function can be calculated only approximately. If $\Theta_r/T << 1$, in fact it is a high temperature approximation, the sum in Eq. (14.1) can be approximated by an integral (take $x = l(l+1)$),

$$q^{(rot)} \approx \int_0^{\infty} e^{-x\Theta_r/T} dx = \frac{T}{\Theta_r} = \frac{8\pi^2 I k_B T}{h^2}. \tag{14.2}$$

This result is the classical mechanical approximation, as expected. It is worth noting that for homonuclear diatomic molecule, the classical rotational partition function becomes, $q^{(rot)} = \Theta_r/2T$, where the factor of 2 accounts for the degeneracy of the molecule with respect to the interchange of indistinguishable atoms. However, the quantization of the rotational energy levels for linear molecules even with small momentum of inertia is not sufficiently fine for a summation replaced by integral.

Wider applicability has the result in the form of the expansion into series in terms of Θ_r/T (if $\Theta_r/T < 1$), it is a so-called Mulholland partition function,

$$q^{(rot)} \approx (T/\Theta_r)\left\{1 + \frac{1}{3}\left(\frac{\Theta_r}{T}\right) + \frac{1}{15}\left(\frac{\Theta_r}{T}\right)^2 + \frac{4}{315}\left(\frac{\Theta_r}{T}\right)^3 + \cdots\right\}. \tag{14.3}$$

This approximation leads to quite accurate results as witnessed in the scientific literature, see entry "Mulholland partition function" on internet. On the basis of the results obtained from spectroscopy experiments, one can conclude that the classical mechanical approximation is accurate for most molecules except at rather low temperatures. The worst case is the hydrogen molecule which is characterized by $\Theta_r = 85.4K$. For this molecule at

room temperature, the quantum effects yield about 10 percent correction to the classical mechanical partition function.

The quantum mechanical vibrational partition function can be evaluated straightforwardly. For diatomic molecule it is,

$$q^{(vib)} = e^{-u_0/k_BT}\sum_{\nu=0}^{\infty}\exp\left\{-\frac{\hbar\omega}{k_BT}(\nu+\frac{1}{2})\right\}$$
$$= \frac{e^{-(u_0+\frac{1}{2}\hbar\omega)/k_BT}}{[1-\exp(-\frac{\hbar\omega}{k_BT})]} = e^{-u_0/k_BT}\frac{\exp(-\frac{\Theta_v}{2T})}{\left[1-\exp(-\frac{\Theta_v}{T})\right]}, \tag{14.4}$$

where the geometric series $\sum_{\nu=0}^{\infty}x^{\nu} = 1/(1-x)$, was used. Here, Θ_v is the characteristic vibrational energy in temperature units,

$$\Theta_v \equiv \hbar\omega/k_B. \tag{14.5}$$

In the high temperature approximation, $\Theta_v/T << 1$, the result above leads to the limiting value,

$$q^{(vib)} \approx e^{-u_0/k_BT}\exp\left(-\frac{\Theta_v}{2T}\right)\left(\frac{T}{\Theta_v}\right). \tag{14.6}$$

This expression differs from the classical mchanical result in that the energy term $u_0 + \frac{1}{2}\hbar\omega$ replaces the classical reference energy u_0 in the partition function. This discrepancy arises because the lowest quantum mechanical energy of a harmonic oscillator is $\hbar\omega/2$. As a consequence, the classical mechanical expression in Eq. (13.42) is not adequate, because within the classical consideration there are no means to include the lowest quantum state energy.

To conclude, we refer to the Table 14.1 that gives values of parameters Θ_r, Θ_v, as well as r_0 and u_0 for different molecules, all these data follow from spectroscopic measurements. Inspecting the numbers in the Table 14.1, one can make conclusions about the range of temperatures within which approximations given in the previous and this chapter are valid. It becomes evident that rotational motion can be approximated by the classical mechanical result under several circumstances. In contrast, vibrational motion requires quantum mechanical treatment, unless very high temperatures are considered. In Table 14.1 force constants for several diatomic molecules are presented.

Table 14.1: Parameters for some diatomic molecules*.

Molecule	Θ_v (K)	Θ_r (K)	u_0 (eV)
H_2	6210	85.300	4.454
O_2	2256	2.070	5.080
NO	2719	2.450	5.290
HBr	3787	12.020	3.600
Cl_2	808	0.351	2.480
HCl	4140	51.2	4.43

*These parameters were obtained from a variety of sources.

Now, we would like to provide expressions describing the contribution of vibrations into thermodynamic properties of a gas of polyatomic molecules, using quantum mechanical result for $q^{(vib)}$. The free energy, following from Eq. (14.4) is,

$$A^{(vib)} = N\left\{u_0 + \frac{1}{2}k_B\Theta_v + k_BT\ln(1-e^{-\Theta_v/T})\right\}. \tag{14.7}$$

Then, the entropy and the internal energy follow from thermodynamic relations,

$$S^{(vib)} = Nk_B\left\{-\ln(1-e^{-\Theta_v/T}) + \frac{\Theta_v/T}{e^{\Theta_v/T}-1}\right\}, \tag{14.8}$$

and

$$U^{(vib)} = N\left\{u_0 + \frac{1}{2}k_B\Theta_v + \frac{k_B\Theta_v}{e^{\Theta_v/T}-1}\right\}. \tag{14.9}$$

Finally, the constant volume heat capacity that comes from vibrations is,

$$C_V^{(vib)}/Nk_B = \frac{(\Theta_v/T)^2 e^{\Theta_v/T}}{(e^{\Theta_v/T}-1)^2}. \tag{14.10}$$

This result requires analysis of its temperature dependence. In the low temperature limit, i.e. $T/\Theta_v \to 0$ or $\Theta_v/T \to \infty$, $C_V^{(vib)}/Nk_B$ tends to zero as $(\Theta_v/T)^2 exp^(-\Theta_v/T)$. On the other hand, in the high temperature limit $T/\Theta_v \to \infty$ or $\Theta_v/T \to 0$, $C_V^{(vib)}/Nk_B \to 1$.

One can combine these conclusions with the results for constant volume heat capacity coming from translational motion and from rotational motion, described in the last paragraph of chapter 13. Thus, one would expect to obtain plateaus corresponding to 3/2, 5/2 and 7/2 at certain temperature intervals for C_V/Nk_B (for the total heat capacity). This is precisely, what the experimental observations provide, see the Fig. 14.1 below.

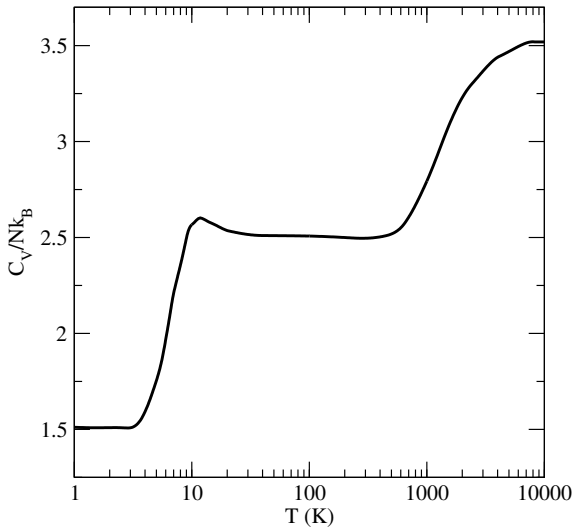


Fig. 14.1: Heat capacity at constant volume of a dilute gas of HCl predicted using spectroscopic measurements. The characteristic rotational and vibrational temperatures, Θ_r and Θ_v are taken from Table 14.1. Reproduced from Ref. [2].

The description of vibrational motion of molecules is intrinsically required for the modelling of solids and the study of thermodynamic properties of this kind of systems.

14.2 Elements of statistical thermodynamics of ideal solids

Let us begin with the precise definition of the system of our study in this subsection. Define an ideal solid as a system of N particles that are assigned to the vicinity of sites fixed in space at coordinates $\mathbf{R}_1,...,\mathbf{R}_N$, The particles can be identical or different, but are distinguishable because they are attributed to the locations of particular lattice sites. The potential energy of the system $U_N(\mathbf{r}_1,...,\mathbf{r}_N)$ attains minimum when all the particles are at the coordinates of lattice sites, i.e. $U_N(\mathbf{r}_1,...,\mathbf{r}_N) > U_N(\mathbf{R}_1,...,\mathbf{R}_N)$. The potential energy of this latter configuration of particles $U_N(\mathbf{R}_1,...,\mathbf{R}_N)$, is denoted as $U_N^{(0)}$. Moreover, it is assumed that the particles are sufficiently tightly bound to their respective sites, such that one can apply expansion into Taylor series about the minimum energy configuration and without loss of accuracy truncate the series after second order term in terms of the displacements $\mathbf{r}_i - \mathbf{R}_i$ of particles from the coordinates of lattice sites.

Let $\xi_1,...,\xi_{3N}$ denote the Cartesian coordinates of particles relative to the lattice sites, i.e. $\xi = r_{1x} - R_{1x},...,\xi_{3N} = r_{Nz} - R_{Nz}$. So, the starting point of our consideration is the classical Hamiltonian,

$$H = \sum_{i=1}^{3N} \frac{1}{2} m_i \dot{\xi}_i^2 + \frac{1}{2} \sum_{i,j}^{3N} a_{ij} \xi_i \xi_j + U_N^{(0)}, \tag{14.11}$$

where the first term yields the kinetic energy, m_i is the mass associated with the coordinate, ξ_i, and,

$$a_{ij} = \left(\frac{\partial^2 U_N}{\partial \xi_j \partial \xi_j} \right) \Big|_{\mathbf{r}^N = \mathbf{R}^N}. \tag{14.12}$$

The matrix with the elements a_{ij} we deal with is symmetric, i.e. $a_{ij} = a_{ji}$. There exists an orthogonal transformation (linear transformation) of the initial coordinates, $\xi_1,...,\xi_{3N}$, into another set of coordinates $\zeta_1,...,\zeta_{3N}$,

$$\zeta_i = \sum_{i=j}^{3N} \alpha_{ij} \xi_j, \tag{14.13}$$

in terms of which the starting Hamiltonian, or better say the potential energy term of the Hamiltonian, becomes diagonal, i.e.

$$H = \sum_{i=1}^{3N} \frac{1}{2} \left[m \dot{\zeta}_i^2 + \lambda_i \zeta_i^2 \right] + U_N^{(0)}, \tag{14.14}$$

where m is certain reference atom mass. So what we are dealing with is a set of oscillators described classically.

Comment: A more detailed description of the diagonalization procedure is given in the supplement to this Lecture.

The term describing the kinetic energy in Eq. (14.14) can be written in the form familiar for us by using the definition of momenta, $p_i = m\dot{\zeta}_i$. Then, the classical partition function can be evaluated,

$$\begin{aligned} Q_N &= e^{-U_N^{(0)}/k_B T} \frac{1}{h^{3N}} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} dp_1 ... dp_{3N} d\zeta_1 ... d\zeta_{3N} \exp\left\{ -\sum_{i=1}^{3N} \left(\frac{p_i^2}{2m} + \frac{1}{2} \lambda_i \zeta_i^2 \right) / k_B T \right\} \\ &= \left(\frac{2\pi k_B T}{h} \right)^{3N} \prod_{i=1}^{3N} \left(\frac{m}{\lambda_i} \right)^{1/2} e^{-U_N^{(0)}/k_B T}. \end{aligned} \tag{14.15}$$

Then, the Helmholtz free energy is,

$$A = -3Nk_B T \ln\left(\frac{2\pi k_B T}{h} \right) - \frac{1}{2} k_B T \sum_{i=1}^{3N} \ln\left(\frac{m}{\lambda_i} \right) + U_N^{(0)}. \tag{14.16}$$

The expressions for entropy and for internal energy follow straightforwardly, as we have done in several places above. Finally, the constant volume heat capacity can be obtained,

$$C_V = 3Nk_B. \tag{14.17}$$

This equation corresponds to the classical approximation for the free energy and is known as the the law of Dulong and Petit. It is an approximation and is valid only above certain temperatures (better say high temperatures) and for certain substances. However, it is known from experiments that all solids exhibit a decreasing heat capacity with decreasing temperature. Thus one can conclude that predictions of the classical mechanical approach for heat capacity are not satisfactory and as expected one should resort to quantum

mechanical treatment. Let us consider a version of quantum mechanical considerations and write down the Schrödinger equation for a solid in the form,

$$\sum_{i=1}^{3N} H_i\psi + U_N^{(0)}\psi = E\psi, \tag{14.18}$$

for the Hamiltonian,

$$H_i = -\frac{\hbar^2}{2m_i}\frac{\partial^2}{\partial\zeta_i^2} + \frac{1}{2}\lambda_i\zeta_i^2. \tag{14.19}$$

From the solution of Schrödinger equation one obtains the energy levels for each of the $3N$ independent harmonic oscillators,

$$E_{i\nu} = \hbar\omega_i\left(\nu + \frac{1}{2}\right), \quad \nu = 0,1,2,..., \tag{14.20}$$

where $\omega_i = (\lambda_i/m)^{1/2}$. The partition function then is (recall that oscillators of the ideal solid are assigned to distinguishable sites),

$$Q_N = e^{-U_N^{(0)}/k_BT}\prod_{i=1}^{3N} q_i, \tag{14.21}$$

$$q_i = \sum_{\nu=0}^{\infty}\exp\left\{-\hbar\omega_i\left(\nu + \frac{1}{2}\right)/k_BT\right\} = \frac{e^{-\hbar\omega_i/2k_BT}}{(1-e^{-\hbar\omega_i/k_BT})}. \tag{14.22}$$

Thermodynamic properties then follow straightforwardly. Again, one should write down the partition functions, derive the Helmholtz free energy, entropy and get the internal energy. Specifically, we are interested in the constant volume specific heat of the solid,

$$C_V = \sum_{i=1}^{3N}\frac{[(\hbar\omega_i)^2/k_BT^2]e^{-\hbar\omega_i/k_BT}}{(1-e^{-\hbar\omega_i/k_BT})^2}. \tag{14.23}$$

Concerning the application of this result we observe that is is necessary to have the frequencies of oscillators, ω_i. This is equivalent to the calculation of the eigenvalues of the matrix $\mathbf{A}$ (see Eq. (14.12) transformed into Eq. (14.14)). The eigenvalues depend on the type of lattice under study as well as on the parameters m_i and force constants a_{ij}. In general, the eigenvalue problem cannot be solved because of size $3N$ and other complexities. The theory developed by Einstein circumvent this difficulty by assuming that all the eigenvalues of the matrix $\mathbf{A}$ are equal. In other words, it is assumed that all possible modes of oscillations are characterized by the same frequency ω_E. Then, the result for heat capacity above, can be written in a simpler and more transparent form,

$$C_V = 3Nk_B\left(\frac{\Theta_E}{T}\right)^2\frac{e^{-\Theta_E/T}}{(1-e^{-\Theta_E/T})^2}, \tag{14.24}$$

where $\Theta_E \equiv \hbar\omega_E/k_B$. In the high temperature limit, $\Theta_E/T \to 0$, one recovers the classical mechanical result, $3Nk_B$. However, the quantum mechanical procedure provides the function that depends on temperature. Moreover, the theory of Einstein is in reasonable agreement with experimental data as it follows from the Fig. 14.2.

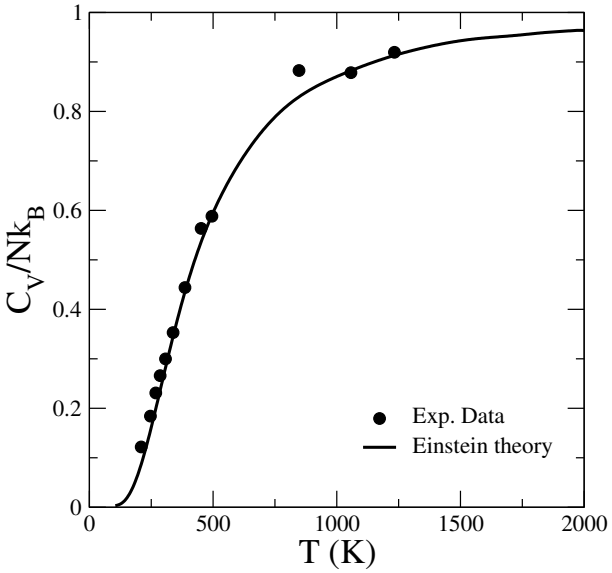


Fig. 14.2: Constant volume heat capacity of diamond as a function of the temperature. The experimental data are compared to Einstein's theory with $\Theta_E = 1320$ K.

15. Paramagnetic materials

15.1 Comments on magnetic materials

General comments

Paramagnetism is a form of magnetism whereby some materials are weakly ”attracted” by an externally applied magnetic field, and form internal, induced magnetic fields in the direction of the applied magnetic field. In contrast to this behavior, diamagnetic materials are ”repelled” by magnetic fields and form induced magnetic fields in the direction opposite to that of the applied magnetic field. Diamagnetism is a quantum mechanical effect that occurs in all materials. If it is the only contribution to the magnetism, the material is called diamagnetic.

However, other forms of magnetism (such as ferromagnetism or paramagnetism) are so much stronger, such that when multiple different forms of magnetism are present in a material, the diamagnetic contribution is usually negligible.

In paramagnetic and ferromagnetic substances, the weak diamagnetic ”force” is overcome by the attractive force of magnetic dipoles in the material. The magnetic permeability of diamagnetic materials is less than the permeability of vacuum, μ_0. In most materials, diamagnetism is a weak effect which can only be detected by sensitive laboratory instruments, Interestingly, a superconductor acts as a strong diamagnetic and can be called as perfect diamante, because it repels a magnetic field entirely from its interior.

Paramagnetic materials include most chemical elements and some compounds; they have a relative magnetic permeability slightly greater than unity (i.e., a small positive magnetic susceptibility) and hence are attracted to magnetic fields. The magnetic moment induced by the applied field is linear in the field strength and rather weak.

Paramagnetism is due to the presence of unpaired electrons in the material, so most atoms with incompletely filled atomic orbitals are paramagnetic, although exceptions such as copper exist. Due to their spin, unpaired electrons have a magnetic dipole moment and act like minuscule magnets. An external magnetic field causes the electrons' spins to align parallel to the field, causing a net attraction. Paramagnetic materials include aluminium, oxygen, titanium, and iron oxide.

Constituent atoms or molecules of paramagnetic materials have permanent magnetic moments (dipoles), even in the absence of an applied field. The permanent moment generally is due to the spin of unpaired electrons in atomic or molecular electron orbitals. In pure paramagnetism, the dipoles do not interact with one another (like ”ideal gas”) and

are randomly oriented in the absence of an external field due to thermal motion, resulting in zero net magnetic moment. When a magnetic field is applied, the dipoles will tend to align with the applied field, resulting in a net magnetic moment in the direction of the applied field. In the classical description, this alignment can be understood to occur due to a torque being provided on the magnetic moments by an applied field, which tries to align the dipoles parallel to the applied field. However, the true origins of the alignment can only be understood via the quantum mechanical properties of spin and angular momentum.

If there is sufficient energy exchange between neighbouring dipoles, they will interact, and may *spontaneously* align or anti-align and form magnetic domains, resulting in ferromagnetism or antiferromagnetism, respectively. Paramagnetic behavior can also be observed in ferromagnetic materials that are above their so-called Curie temperature, and in antiferromagnets above their Neel characteristic temperature. At these temperatures, the thermal motion effect overcomes the interaction energy between the spins.

In general, paramagnetic effects are quite small: the magnetic susceptibility is of the order of 10^{-3} to 10^{-5} for most paramagnets, but may be as high as 10^{-1} for synthetic paramagnets such as ferrofluids.

Paramagnetism is most easily observed in the salts of some of the first row transition metals (manganese through nickel). These metal ions have unpaired electrons in degenerate d orbitals and thus exhibit paramagnetism. But strong-field ligands can split the energy levels of the d orbitals so that they are no longer degenerate. Salts of these ions that have strong-field ligands are diamagnetic.

Water is not attracted to the magnet poles, it is diamagnetic. However, the interaction is extremely weak, and visually noticeable effects can only be obtained using powerful superconducting magnets. In the past, demonstrations of diamagnetism in water using more accessible permanent magnets have been qualitative only. If a powerful magnet (such as a supermagnet) is covered with a layer of water (that is thin compared to the diameter of the magnet) then the field of the magnet significantly repels the water. This causes a slight concave dimple in the water's surface that may be seen by a reflection in its surface. On the other hand, there are no unpaired electrons in sodium chloride, it is diamagnetic as well.

Other relevant examples are the following: Manganese(II) sulfate monohydrate is strongly attracted by the magnet, which shows that it is paramagnetic. Manganese ions in manganese(II) sulfate monohydrate have five unpaired electrons. Iron(II) sulfate pentahydrate can be brought closer to the magnet than the manganese(II) sulfate monohydrate, but eventually it swings toward the magnet. This compound is not as strongly paramagnetic as the manganese(II) sulfate. Iron(II) ions have four unpaired electrons. Cobalt(II) chloride hexahydrate can be brought closer to the poles than the iron (II) sulfate or the manganese(II) sulfate before it swings toward the magnet. Cobalt(II) ions have three unpaired electrons. When nickel(II) sulfate hexahydrate is brought near the poles of the magnet, it appears to have about the same attraction as the cobalt(II) chloride, perhaps slightly less. Nickel(II) ions have two unpaired electrons.

15.2 Statement and solution of the problem

We consider a system of N identical atoms with magnetic moment that are independent each other. In fact, we are interested in the paramagnet consisting of a set of atoms in a crystal lattice. Because the atoms are situated at particular positions we can tell which is which - that is to say they are distinguishable because they are localized. The system is subjected to an external magnetic field $\mathbf{H}$. The magnetic energy of an atom is,

$$\varepsilon = -\boldsymbol{\mu}\mathbf{H}, \tag{15.1}$$

where $\boldsymbol{\mu}$ is the magnetic moment of an atom.

The magnetic moment of an atom is proportional to its angular momentum $k_B\mathbf{J}$, as follows,

$$\boldsymbol{\mu} = g\mu_0\mathbf{J}, \tag{15.2}$$

where g is the g-factor and μ_0 denotes the Bohr magneton, $\mu_0 = e\hbar/2mc$ ($\mathbf{J}$ is dimensionless). If the magnetic field $\mathbf{H}$ is directed along the z- axis, the equations above become,

$$\varepsilon = -g\mu_0 J_z H, \quad \mu_z = g\mu_0 J_z, \tag{15.3}$$

where according to the laws of quantum mechanics, the z component of the angular momentum vector takes values,

$$J_z = -J, -J+1, ..., .., J-1, J. \tag{15.4}$$

The characteristic spin moment J is integer or half integer. If an external magnetic field is applied, the degeneracy is eliminated. Schematic representation of the setup for the two-state system is presented in Fig. 15.1.

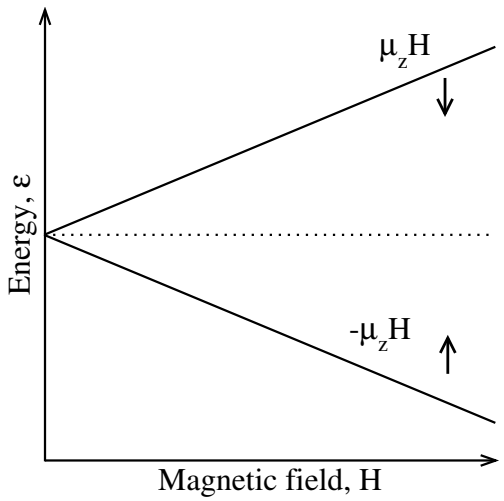


Fig. 15.1: Schematic representation of the setup for two-state system.

The single particle partition function of the canonical ensemble that describes magnetic properties, with energy levels according to Eq. (15.3) is,

$$q_m = \sum_{m=-J}^{J} \exp\{g\mu_0 mH/k_B T\}. \tag{15.5}$$

It is convenient to introduce notation $\zeta = g\mu_0 H/k_B T$. Using the properties of geometric series, it is not difficult to obtain the result for the partition function in the form,

$$q_m = \frac{e^{-\zeta J} - e^{\zeta(J+1)}}{1 - e^{\zeta}}. \tag{15.6}$$

From, this result, if $J = S = 1/2$, $L = 0$ ($J = S + L$), and the definition [23],

$$g = \frac{3J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}, \tag{15.7}$$

we observe that $g = 2$, and then one can arrive at the partition function for the particular case of the two-state, $-1/2, 1/2$, system,

$$q_m = 2\cosh\left(\frac{\zeta}{2}\right) = 2\cosh\left(\frac{\mu_0 H}{k_B T}\right). \tag{15.8}$$

Let us elaborate this simple case first. The internal energy of the system of N independent momenta follows from common, and previously used, relation, $U = -N\partial \ln q_m/\partial\beta$ ($\beta = 1/k_BT$),

$$U = -N\mu_0 H \tanh\left(\frac{\mu_0 H}{k_B T}\right). \tag{15.9}$$

Schematic dependence of internal energy on temperature and on the strength of magnetic field is shown in panels a and b of Fig. 15.2.

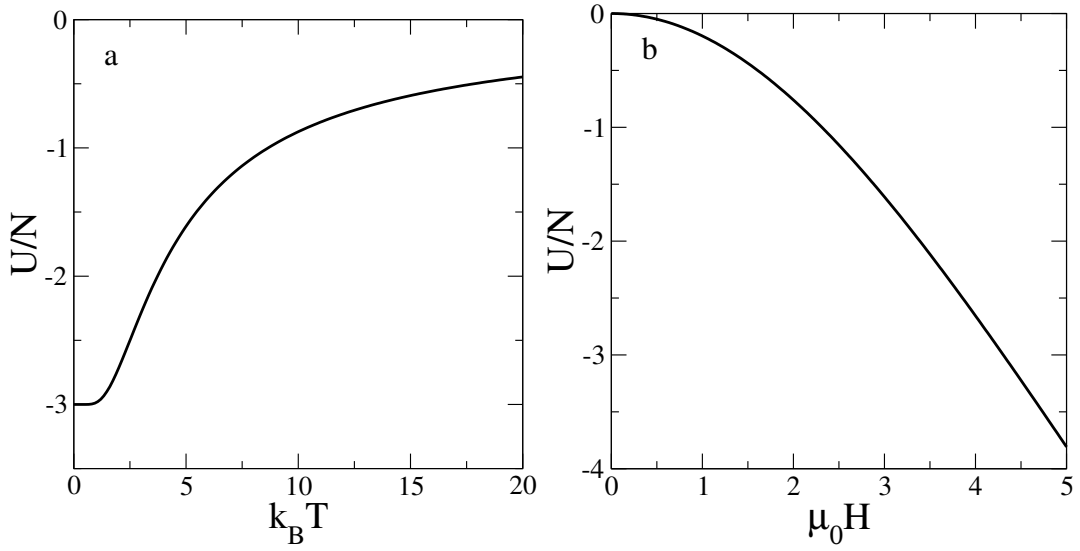


Fig. 15.2: The internal energy per spin as a function of temperature (panel a) and of the strength of the magnetic field (panel b). In panel a, $\mu_0 H = 3$ and in panel b, $k_BT = 5$.

The magnetic contribution to the constant volume heat capacity follows from the internal energy result straightforwardly,

$$C_{V,m} = \frac{dU}{dT} = Nk_B\left(\frac{\mu_0 H}{k_B T}\right)^2 \frac{1}{\cosh^2(\mu_0 H/k_B T)} = Nk_B\left(\frac{\theta}{T}\right)^2 \frac{e^{\theta/T}}{(e^{\theta/T}+1)^2}, \tag{15.10}$$

where the notation $\theta = 2\mu_0 H/k_B$ has been used. This permits to perform analyses of the limiting cases. Namely,

$$C_{V,m} \propto 1/T^2, if\ T \to \infty; \quad C_{V,m} \propto (1/T^2)e^{-\theta/T}, if\ T \to 0. \tag{15.11}$$

The entire curve following from Eq. (15.10) can be plotted and has the form shown in the Fig. 15.3.

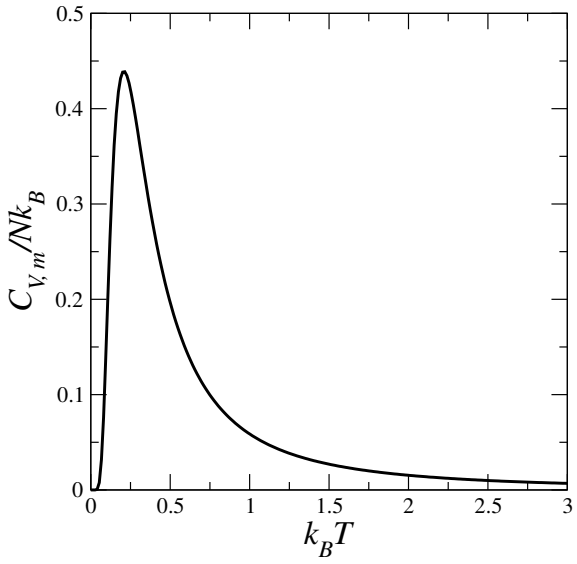


Fig. 15.3: The dependence of heat capacity as a function of the temperature ($\theta = 1$).

We observe that the theory (based on Boltzmann statistics) predicts the so-called Schottky anomaly. It is an effect where the specific heat capacity of a solid at low temperature exhibits a maximum. The location of a maximum depends on the value of applied external magnetic field. It is called anomalous because the heat capacity commonly increases with temperature, or remains constant.

In addition, it is helpful to explore the behavior of entropy. It follows from Eq. (15.8),

$$S = Nk_B \ln\left[2\cosh\left(\frac{\mu_0 H}{k_B T}\right)\right] - (N\mu_0 H/T)\tanh\left(\frac{\mu_0 H}{k_B T}\right). \tag{15.12}$$

Two limits should be explored. In the high temperature limit, $k_B T >> \mu_0 H$, the entropy tends to,

$$S \to Nk_B \ln(2), \tag{15.13}$$

whereas in the limit of low temperature, $k_B T << \mu_0 H$, the entropy vanishes, $S \to 0$. Both results are physically well justified. At very low temperature all the electrons should be in the state with lower energy, therefore entropy vanishes. On the other hand, at very high temperatures, the thermal energy $k_B T$ is much higher than the magnetic splitting between energy levels, thus two magnetic energy states are equiprobable.

Finally, it is necessary to comment on the magnetization, M, of the system of magnetic moments. The magnetization of N independent magnetic moments is,

$$M = N < \mu_z >, \tag{15.14}$$

where $< \mu_z >$ is the average magnetic moment for an atom calculated with the distribution function of the canonical ensemble,

$$< \mu_z >= \sum_{m=-J}^{J} g\,\mu_0 m\,\exp\left(g\mu_0 mH/k_B T\right)/q_m = k_B T\left(\frac{\partial \ln q_m}{\partial H}\right)\Big|_T. \tag{15.15}$$

Then, the magnetization of the system of N independent magnetic moments, with the free energy $A = -Nk_B T \ln q_m = U - TS$, in Eq. (15.14) is,

$$M = Nk_B T\left(\frac{\partial \ln q_m}{\partial H}\right)\Big|_T = -\left(\frac{\partial A}{\partial H}\right)\Big|_T = N\mu_0\,\tanh\left(\frac{\mu_0 H}{k_B T}\right). \tag{15.16}$$

This result permits to obtain the so-called Curie law for paramagnetism, describing the dependence of magnetization on the applied magnetic field in the region of weak fields,

$$M = \frac{N\mu_0^2 H}{k_B T}. \tag{15.17}$$

This is nothing else than the equation of state of paramagnetic system in terms of M, H, and T variables. In Fig. (15.4), the magnetization M is plotted as a function of the dimensionless parameter $\mu_0 H/k_B T$, illustrating the thermal behavior of the magnetic response of the system upon the action of the external magnetic field.

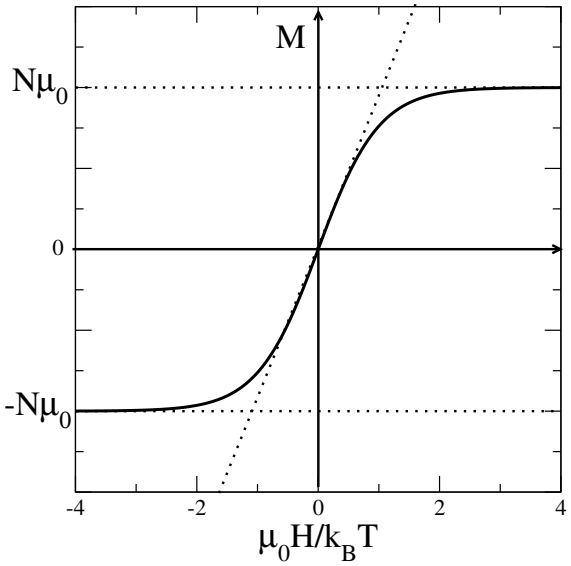


Fig. 15.4: The dependence of the total magnetic moment of the paramagnet, i.e. the magnetization on the strength of magnetic field. The curve is obtained using Eq. (15.16).

A set of more general expressions follows from the single-particle partition function, q_m, given by Eq. (15.6). It refers to arbitrary values of J. Specifically, we can obtain the average magnetic moment of a particle in the paramagnetic system under the influence of external magnetic field,

$$< \mu_z >= \frac{1}{q_m} \sum_{m=-J}^{J} (g\mu_0 m)\, e^{g\mu_0 mH/k_B T}. \tag{15.18}$$

This step is equivalent to what we have done in Eq. (15.9),

$$< \mu_z >= k_B T \left(\frac{\partial \ln q_m}{\partial H} \right)\Big|_T = \frac{g\mu_0}{q_m} \frac{\partial}{\partial \zeta} \left[\frac{(e^{-\zeta J} - e^{-\zeta(J+1)})}{(1 - e^{\zeta})} \right], \tag{15.19}$$

where $\zeta = H/T$. A set of experimental data illustrates what we are looking for within the theoretical procedure.

The derivative in Eq. (15.19) can be taken explicitly to arrive at average magnetization of a sample of N magnetic moments per volume,

$$\overline{M} = \frac{N}{V} < \mu_z >= \frac{N}{V} g\mu_0 J B_J(\zeta). \tag{15.20}$$

This final result involves the so-called Brillouin special function, $B_J(x)$,

$$B_J(x) = \frac{1}{J} \left\{ (J + \frac{1}{2}) \coth[(J + \frac{1}{2})x] - \frac{1}{2} \coth(\frac{1}{2}x) \right\}. \tag{15.21}$$

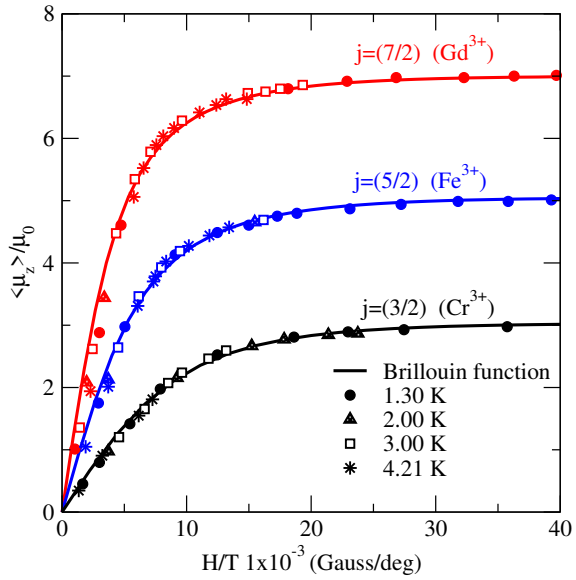


Fig. 15.5: Average magnetization of a sample of a paramagnetic substance as a function of H/T. Experimental data refer to ions of chromium, potassium alum, iron ammonium alum, and gadolinium sulfate octahydrate for which $g = 2$ and $j = 3/2$, $5/2$, and $7/2$, respectively, see Ref. [24].

We can now explore the result of Eq. (15.20) to define the magnetic susceptibility, χ, as follows,

$$\overline{M} = \chi H, \tag{15.22}$$

that describes magnetic response of the sample under study to the action of external magnetic field,

$$\chi = \frac{N}{V}\frac{g\mu_0 J}{H} B_J(\zeta). \tag{15.23}$$

In order to perform analysis of the behavior of magnetic susceptibility we need to comment on the behavior of the Brillouin function. The function $B_J(\zeta)$ is a monotonously increasing function, it equals to zero as $\zeta \to 0$ (it is necessary to apply series expansion of the hyperbolic cotangent to extract this limit). On the other hand, the function $B_J(\zeta)$ tends to unity in the limit $\zeta \to \infty$.

More specifically, in the limit of high temperature or small values of external field, H (this limit has been formulated above, $k_B T >> g\mu_0 H$, to arrive at Eqs. (15.13) and (15.17)), the Brillouin function behaves as $B_J(\zeta) \approx (J+1)\zeta/3$. Then, the susceptibility χ can be written straightforwardly as,

$$\chi = \frac{N}{V}\frac{g^2\mu_0^2 J(J+1)}{3k_B T}. \tag{15.24}$$

This expression just represents another expression for the $1/T$ behavior - in essence the Curie law, obtained already within the analyses given slightly above.

On the other hand, in the limit of strong external fields, H, or at low temperatures, T (recall our discussion of the case $k_B T << g\mu_0 H$ above), it follows from Eq. (15.22) that,

$$\overline{M} = \frac{N}{V} g\mu_0 J, \tag{15.25}$$

or in other words magnetization saturates (or does not depend on the applied field). This physical situation corresponds to alignment of all magnetic moments along the direction of magnetic field.

To summarize the general case, we would like just to give expressions for the magnetic internal energy,

$$U = -Ng\mu_0 J B_J(\zeta) H, \tag{15.26}$$

and for the magnetic contribution to the constant volume specific heat,

$$C_{V,m} = \frac{dU}{dT} = -Nk_B J\zeta^2 \frac{dB_J(\zeta)}{d\zeta}. \tag{15.27}$$

Analysis of the behavior of these expressions on temperature and on the strength of external magnetic field coincides with the discussion given above for the particular case $J = 1/2$, we refer to Eqs. (15.8)-(15.10) of this chapter. The derivative necessary to apply in Eq. (15.26) follows from the definition of the Brillouin function defined in Eq. (15.21),

$$\frac{dB_J(\zeta)}{d\zeta} = -\frac{1}{J}\left\{\frac{(J+1/2)^2}{\sinh^2[(J+1/2)\zeta]} - \frac{1}{4}\frac{1}{\sinh^2(\zeta/2)}\right\}. \tag{15.28}$$

Several examples concerning the properties of paramagnetic systems that follow from application of Boltzmann statistics can be found in specialized literature [24, 25].

15.3 One-dimensional Ising model

We proceed now to the models of another type of magnetic systems. Namely, let us consider a one-dimensional (1D) Ising model. It is a chain of equidistantly separated N spins that take values ± 1 ($s_i = +1$ or $s_i = -1$, in common scientific slang spin up and spin down). The length of the system is L. The model assumes *interaction* between nearest neighbors only, such that the Hamiltonian of the system is,

$$E_s = -J\sum_{i=1}^{N-1} s_i s_{i+1}. \tag{15.29}$$

The parameter J providing measure of the strength of interaction between nearest spins is of utmost importance. If J is positive, one can see that the nearest spins alignment ($s_i = s_{i+1}$) is favored due to a lower value of energy for such configuration, comparing to oppositely directed spins. Then, the model can be thought as a 1D most simple or primitive model for ferromagnetic, promoting magnetism via ordering of spins.

On the other hand, if J is negative, antialignment of neighboring spins is energetically preferred, i.e. $s_i = -s_{i+1}$. Thus, the model can be thought as related to antiferromagnetism in most simple manner. Intuitively, one should expect that when no external field is applied, the antiferromagnetic structure corresponds to a vanishing total magnetization.

Low dimensionality of the model permits to deal with the partition function,

$$Q_N = \sum_{s_1=-1}^{1}\sum_{s_2=-1}^{1}\ldots\sum_{s_N=-1}^{1}\exp\{\beta J\sum_{i=1}^{N-1} s_j s_{i+1}\}, \tag{15.30}$$

easily. It can be evaluated step by step, e.g.

$$Q_2 = 2\sum_{s_1=-1}^{1}\sum_{s_2=-1}^{1} e^{\beta J s_1 s_2} = \sum_{s_1=-1}^{1}\left(e^{-\beta J s_1} + e^{\beta J s_1}\right) = 2\left(2\cosh(\beta J)\right), \tag{15.31}$$

next Q_3, Q_4,.., to observe that,

$$Q_N = 2(2\cosh(\beta J))^{N-1}. \tag{15.32}$$

Thermodynamic properties then follow straightforwardly, $(U = -\partial \ln Z_N / \partial \beta)$,

$$\begin{aligned} A =& -k_B T \ln 2 - (N-1) k_B T \ln(2\cosh(\beta J)), \\ U =& -(N-1) J \tanh(\beta J), \\ C_V =& (N-1) k_B (\beta J)^2 (1 - \tanh^2(\beta J)). \end{aligned} \tag{15.33}$$

The most interesting is the temperature dependence of the heat capacity, C_V. In the region of high temperatures, $|\beta J| << 1$, having in mind Taylor series of hyperbolic tangent ($\tanh x = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots$), one arrives at the asymptotic expression,

$$C_V \approx (N-1) \frac{J^2}{k_B T^2}. \tag{15.34}$$

In contrast, in the region of low temperatures, $|\beta J| >> 1$, it follows from Eq. (15.33) that,

$$C_V \approx 4(N-1) \frac{J^2}{k_B T^2} \exp\left(\frac{-2|J|}{k_B T}\right). \tag{15.35}$$

One can recover the entire plot of C_V using Eq. (15.33) as well. It can can be seen in Fig. (15.6) that the heat capacity passes through a maximum at certain temperature interval. This behavior of specific heat is the result of competition between magnetic inter-particle interactions and thermal energy, in contrast to what we studied for paramagnetism in the first part of this chapter.

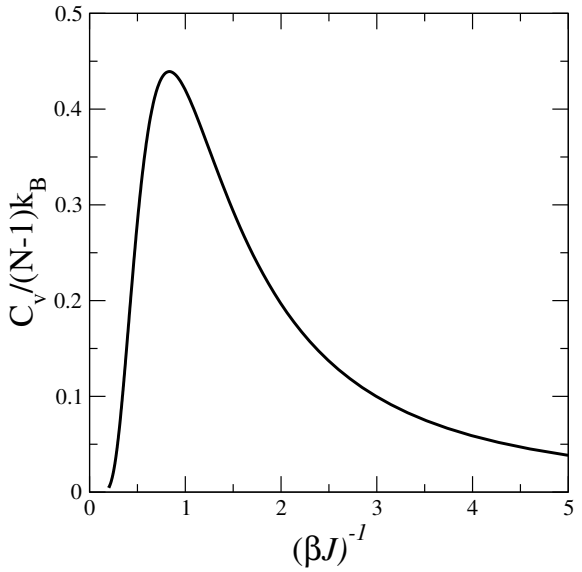


Fig. 15.6: The dependence of the specific heat on temperature for the one-dimensional Ising model.

It is worth mentioning some theoretical predictions concerning heat capacity. Namely, at $T \to 0$ the specific heat vanishes. Under such thermodynamic conditions, the spins obviously enter the lowest energy state, moreover they cannot be excited to the first excited state if $k_B T$ is small comparing to the energy described by the interaction parameter $|J|$. As it has been mentioned in the discussion of the Hamiltonian above, the ground state would be with all spins equally aligned when $J > 0$, i.e. all of them will be in the configuration, $s_i = +1$ or $s_i = -1$. If $J < 0$ the ground state will correspond to all neighboring spins oppositely aligned, $s_i = -s_{i+1}$. In the opposite limit, $T \to \infty$, the spins would attain random configuration (magnetic ordering effects will be negligible comparing to thermal motion effects) the internal energy becomes small and the specific heat vanishes as well.

Interestingly, the specific heat goes through a maximum at some temperature but it is infinitely differentiable w.r.t. temperature, consequently no magnetic transition of the order-disorder type is predicted by the model. Two issues are of importance in this aspect. One is that the interaction between spins is only of the nearest neighbors type (perhaps insufficient to yield transition). The other principal issue is the dimensionality of the model. It has been proven and explained in the book "Statistical Physics" by Landau and Lifshitz [10] that there are no true phase transitions in one-dimensional systems.

Final issue of this chapter is in discussing the effect of external magnetic field on the behavior of 1D Ising model. To do that, one needs to extend the Hamiltonian given by Eq. (15.29) as follows,

$$E_s = -J \sum_{i=1}^{N-1} s_i s_{i+1} - h \sum_{i=1}^{N} s_i, \tag{15.36}$$

where the external field is, $h = \mu H$ (μ is the magnitude of the magnetic spin moment). The evaluation of the partition function is more complex and is omitted in the present text. We just refer to the result for magnetization obtained from the free energy,

$$M = -n\mu \left(\frac{\partial A}{\partial h} \right) \Big|_{\beta} = \frac{n\mu \, \sinh(\beta h)}{[\sinh^2(\beta h) + \exp(-4\beta J)]^{1/2}}, \tag{15.37}$$

where n is the number density of spins ($n = N/L$). Note similarity of the first equality sign above, with our previous calculation given in Eq. (15.20). One can plot the magnetization on the external field according to the result presented, see the Fig. (15.7) below.

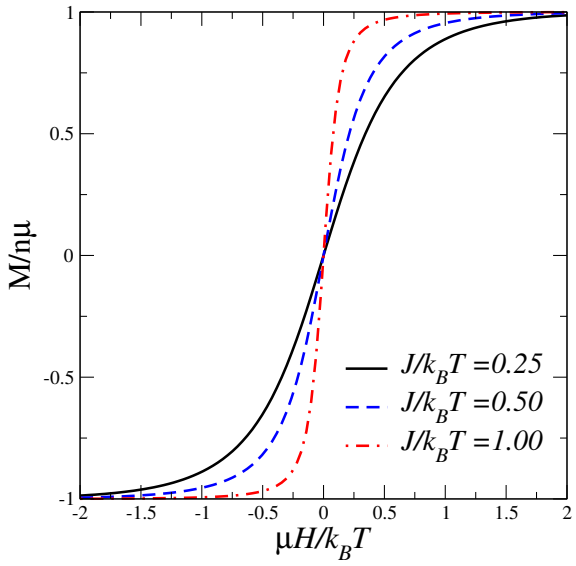
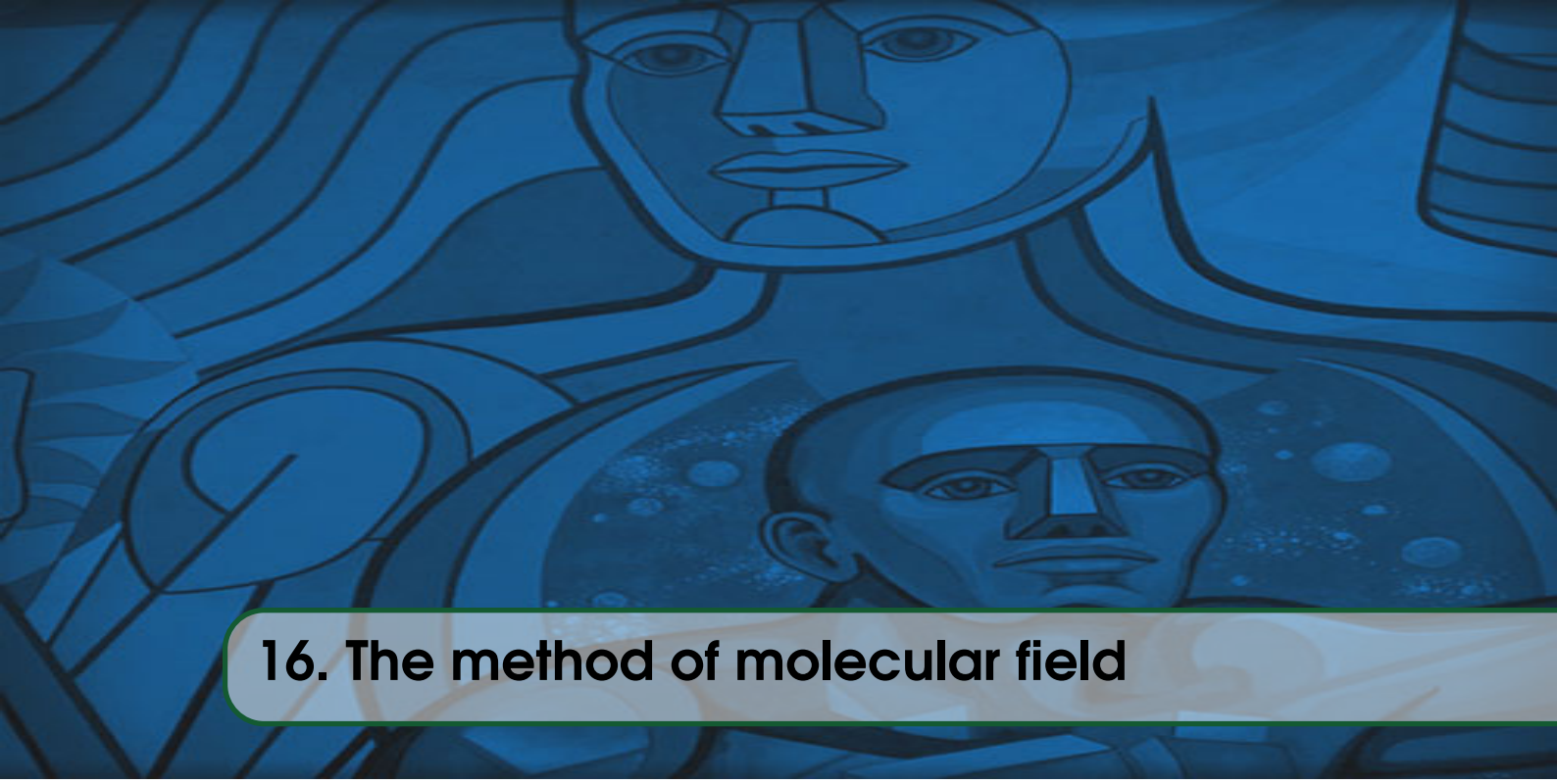


Fig. 15.7: Magnetization isotherms as a function of the external field strength for the one-dimensional Ising model in the magnetic field. Redrawn from T. Davis [2].

In conclusion, the magnetization is a smooth function of the external field strength for all finite temperatures. Thus, the system does not show order-disorder transition or equivalently the field is unable to qualitatively change the behavior of magnetization, appearance of spontaneous magnetization is not captured. This is an illustration of a general rule, if the system does not possess phase transition without external field, the field is unable to generate phase transition on its own.



16. The method of molecular field

16.1 The method of molecular field for ferromagnetism

General comments.

We have seen that only for ideal type systems one can obtain thermodynamic properties quite easily. If the system deviates from ideal type setup, or say becomes non-ideal, then to calculate the partition function becomes problematic. Therefore, the following question arises: is it possible to modify the binary type term describing inter-particle interactions by the function of unary type, something like for the systems under the influence of external field (the field acts on each particle separately) such that thermodynamic properties of an original system and of the one described by a modified Hamiltonian remain close each other. For some systems this trick can be really done.

Statement and solution of the problem.

Similar to the previous chapter consider a system of N spins on a lattice of equivalent sites. It is assumed that the magnetic moment on the site can take only two values $\pm\mu_B$, where μ_B is the Bohr magneton. Therefore, we are dealing with the following Hamiltonian,

$$E = -\frac{1}{2}\sum_{f}\sum_{f',f\neq f'} J_{ff'}s_f s_{f'}, \tag{16.1}$$

where $s_i = \pm 1$, f and $f^{'}$ are the indices of the lattice sites. The summation is over all the sites of the system. The parameter $J_{ff'}$ is chosen positive and depends only on the distance between the sites f and $f^{'}$. However, the interaction is assumed between the nearest neighbor sites only. The energy of the system will be at minimum if all the spins are parallel to a chosen direction, i.e. if all the numbers are either $+1$ or -1. If $J_{ff'} > 0$, the ground state of the Ising model represents ordered ferromagnetic. For the sake of convenience, assume that all spins have orientation -1 in the ground state.

Let us develop the molecular field approximation. To do that, each of spins in the Hamiltonian is substituted by the identity,

$$s_f \equiv < s_f > + s_f - < s_f > \equiv < s_f > + \Delta s_f = < s > + \Delta s_f. \tag{16.2}$$

Here, $< s_f >$ is the statistical average of the value of the variable s_f at a site f. It is important to note that $< s_f > = < s >$ because all the lattice sites are equivalent and the

average does not depend on the site index, whereas the deviation depends on the site index.

Now, substitute this notation into the original Hamiltonian given by Eq. (16.1), to get,

$$E = -\frac{1}{2}\sum_{f\neq f'} J_{ff'} < s >^2 - \sum_{f\neq f'} J_{ff'} < s > \Delta s_{f'} - \frac{1}{2}\sum_{f\neq f'} J_{ff'}\Delta s_f \Delta s_{f'}. \tag{16.3}$$

Essence of the molecular field approximation is in neglecting the last term in the r.h.s. of the equation above. In other words, the squared term for the deviations is neglected. Then, we have,

$$E_M = -\frac{1}{2}\sum_{f\neq f'} J_{ff'} < s >^2 - \sum_{f\neq f'} J_{ff'} < s > \Delta s_{f'}. \tag{16.4}$$

Now, we would like to return to the original variable s_f, by using substitution,

$$\Delta s_{f'} = s_{f'} - < s >, \tag{16.5}$$

according to Eq. (16.2). In addition, we note that,

$$\sum_{f'} J_{ff'} = J_f = J, \tag{16.6}$$

i.e. it does not depend on on the site index, because all the sites are equivalent. In consequence, the Hamiltonian of the Ising model within the molecular field approximation will have the form,

$$E_M = \frac{1}{2}NJ < s >^2 - \sum_f J < s > s_f. \tag{16.7}$$

The dependence on the true spin variables (not on averages) in this Hamiltonian is the same as in the Hamiltonian of a system of independent spins under the influence of external magnetic field, H,

$$E_{id} = -\sum_f \mu_B H s_f, \tag{16.8}$$

cf. the second term of Eq. (15.36). Therefore, we can interpret the term,

$$\frac{J < s >}{\mu_B} = H_M \tag{16.9}$$

as the internal magnetic field which intuitively can be responsible for the appearance of spontaneous magnetic order (frequently it is called as Weiss molecular field). If, in addition, the external magnetic field is present, then the Hamiltonian will have the form,

$$E_M = \frac{1}{2}NJ < s >^2 -(J < s > +\mu_B H)\sum_f s_f. \tag{16.10}$$

This Hamiltonian is the sum of contributions coming from individual atoms and consequently the partition function will be given as a product of one-particle partition functions,

$$Q_M = \prod_{f=1}^{N} q_f, \tag{16.11}$$

with

$$q_f = \sum_{s_f=-1}^{+1} \exp\left\{-\frac{J<s>^2/2-(J<s>+\mu_B H)s_f}{k_B T}\right\}$$
$$= \exp\left\{-\frac{J<s>^2}{2k_B T}\right\} 2\cosh\left(\frac{J<s>+\mu_B H}{k_B T}\right). \tag{16.12}$$

It is worth noting, that the one-particle partition function does not depend on the site index because all the sites are equivalent. The partition function, Q_M, then leads to the free energy,

$$A_M = \frac{1}{2}NJ<s>^2 - Nk_B T \ln\left[2\cosh\left(\frac{J<s>+\mu_B H}{k_B T}\right)\right]. \tag{16.13}$$

One can formally calculate then all thermodynamic properties of the model ferromagnetic in the molecular field approximation. Still, this expression contains unknown parameter $<s>$. Let us calculate it.

First note that magnetization (average magnetization) of the system can be related to the difference of the average numbers of differently oriented spins,

$$M = \mu_B(<N_->-<N_+>). \tag{16.14}$$

On the other hand, by the definition of an average spin at arbitrary site in terms of the population of spins with different orientation, we have,

$$<s>=\frac{<N_->-<N_+>}{N}. \tag{16.15}$$

Thus,

$$M = \mu_B N<s>= M_{max}<s>. \tag{16.16}$$

Here, $M_{max} = \mu_B N$, is the notation for the maximum magnetization of the system corresponding to the configuration when all the spins are parallel. With these preliminaries, one can write down the free energy of the system as follows,

$$A_M = \frac{1}{2}NJ\frac{M^2}{M_{max}^2} - Nk_B T \ln\left[2\cosh\left(\frac{J}{k_B T}\frac{M}{M_{max}}+\frac{\mu_B H}{k_B T}\right)\right]. \tag{16.17}$$

Thermodynamic definition of magnetization, as a conjugate of the external magnetic field, has been used by us previously in several occasions,

$$M = -\left(\frac{\partial A}{\partial H}\right)\Big|_{T,V}, \tag{16.18}$$

It leads to the following equation,

$$M = M_{max}\ \tanh\left(\frac{T_c}{T}\frac{M}{M_{max}}+\frac{\mu_B H}{k_B T}\right), \tag{16.19}$$

where, for the sake of convenience, we have introduced the notation for the temperature scale, $T_c = J/k_B$. The magnetic susceptibility has been defined in Eq. (15.22). At high temperature the hyperbolic tangent in Eq. (16.19) can be expanded into series preserving the linear term only. Then,

$$\frac{M}{M_{max}} = \chi H = \frac{\mu_B H}{k_B(T-T_c)}, \tag{16.20}$$

evidencing that the susceptibility diverges, if the temperature approaches the Curie temperature from above. This dependence on temperature is the essence of the Curie-Weise's law, that describes the magnetic susceptibility, χ, of a ferromagnet in the paramagnetic region above the Curie temperature,

$$\chi = \frac{C}{T - T_c}. \tag{16.21}$$

If the external magnetic field is absent, then the magnetization of the system is the *spontaneous* magnetization, M_s. It can be obtained from the equation above as,

$$\frac{M_s}{M_{max}} = \tanh\left(\frac{T_c}{T}\frac{M_s}{M_{max}}\right). \tag{16.22}$$

The solution can be found simply plotting both sides of this equation. Introduce notation,

$$\frac{T_c}{T}\frac{M_s}{M_{max}} = x, \tag{16.23}$$

and look for the solutions of the equation,

$$\frac{T}{T_c}x = \tanh(x). \tag{16.24}$$

First plot the hyperbolic tangent for the positive values of x. If $T > T_c$ the inclination of the l.h.s. is high, such that the crossing with the hyperbolic tangent occurs solely at the point $x = 0$. Thus, there is no spontaneous magnetization (see the Fig. (16.1) below).

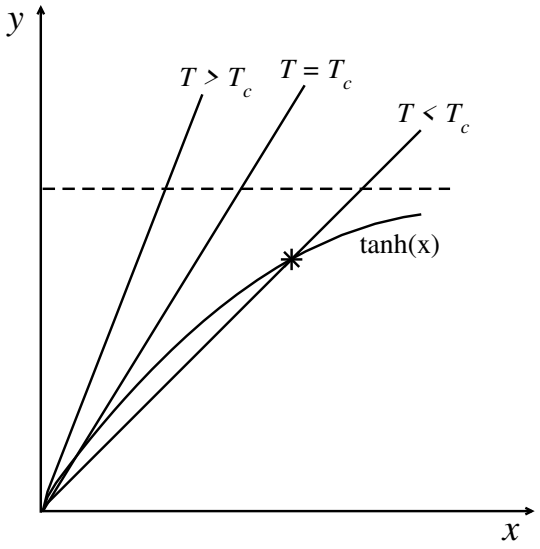


Fig. 16.1: Graphical representation of the solution of the equation for spontaneous magnetization.

If $T < T_c$, there are two crossing points, one - at $x = 0$ is trivial solution whereas another one at $x \neq 0$ is the non-trivial solution. The question is: which of two corresponds to the equilibrium situation? It is known from thermodynamics, that the stable equilibrium for the system with T, V and H independent variables corresponds to the minimum value of the Helmholtz free energy, A. By inspection of the form of the free energy in Eq. (16.17), we see that A_M is lower for the case of non-trivial solution. Thus this solution corresponds to a stable phase and consequently at $T < T_c$ it is characterized by non-zero spontaneous magnetization. The characteristic temperature yielding condition of disappearance of spontaneous magnetization is the so-called Curie temperature.

It is worth to analyze the expression for the spontaneous magnetization for different temperatures. It is easy to make conclusions at low temperature, $T/T_c << 1$ and at high temperature, $T/T_c \approx 1$. At low temperatures, $T/T_c << 1$, the inclination of the l.h.s. of equation above is very low and therefore the spontaneous magnetization, M_s, value is close to the saturation value M_{max}. Consequently,

$$\frac{M_s}{M_{max}}\frac{T_c}{T} >> 1. \tag{16.25}$$

Thus, we have,

$$\frac{M_s}{M_{max}} = \tanh\left(\frac{T_c}{T}\frac{M_s}{M_{max}}\right) = \frac{1-e^{-2(Tc/T)(M_s/M_{max})}}{1+e^{-2(Tc/T)(M_s/M_{max})}} \approx 1-2e^{-2T_c/T}. \tag{16.26}$$

On the other hand, if the temperature is close to Curie temperature, the following inequality holds,

$$\frac{M_s}{M_{max}}\frac{T_c}{T} << 1. \tag{16.27}$$

Then, expanding the hyperbolic tangent into series, we obtain the equation,

$$\frac{M_s}{M_{max}} = \frac{T_c}{T}\frac{M_s}{M_{max}} - \frac{1}{3}\left(\frac{T_c}{T}\frac{M_s}{M_{max}}\right)^3 + \cdots, \tag{16.28}$$

and the result becomes,

$$\left(\frac{M_s}{M_{max}}\right)^2 = 3\left(\frac{T}{T_c}\right)\left(1-\frac{T}{T_c}\right), \tag{16.29}$$

or approximately,

$$\frac{M_s}{M_{max}} = \sqrt{3\frac{T_c-T}{T_c}}. \tag{16.30}$$

The schematic graph of the spontaneous magnetization on temperature is given in the Fig. 16.2. The dash-dotted solid curve describes the result coming from molecular field approximation whereas the solid line roughly corresponds to experimental observations.

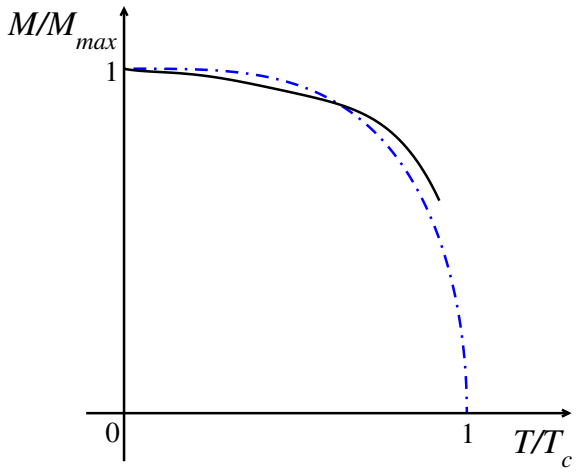


Fig. 16.2: Schematical dependence of the spontaneous magnetization on temperature.

Theoretical developments given above have thermodynamic consequences. Our focus is in the internal energy and constant volume heat capacity. The internal energy results

from the averaging of the Hamiltonian in Eq. (16.10), thus in the framework of molecular field approximation it is,

$$U =< E_M >= -\frac{1}{2}NJ < s >^2 -N\mu_B H < s >= -\frac{1}{2}qM^2 - HM, \tag{16.31}$$

where $q = J/(N\mu_B^2)$. Then, the heat capacity is,

$$C_{V,H} = \left(\frac{dU}{dT}\right)\Big|_{V,H} = -qM\left(\frac{dM}{dT}\right)\Big|_{V,H} - H\left(\frac{dM}{dT}\right)\Big|_{V,H}. \tag{16.32}$$

Because the magnetization (at constant external field value) decreases with increasing temperature (see e.g. the Fig. (16.2) above) the heat capacity is positive ($C_{V,H} > 0$). Moreover, in the absence of external magnetic field,

$$C_V = -qM_s\left(\frac{dM_s}{dT}\right)\Big|_V. \tag{16.33}$$

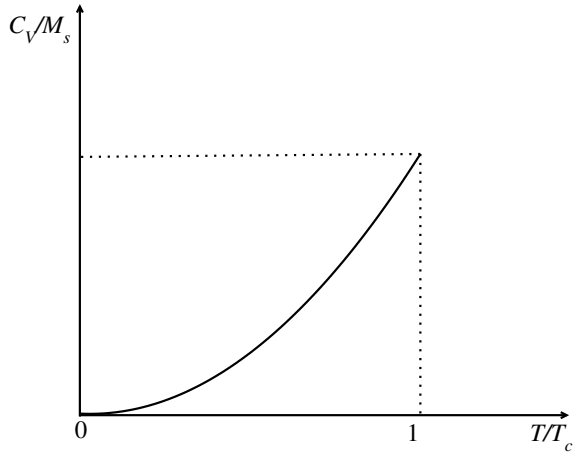


Fig. 16.3: The dependence of heat capacity on temperature.

So, the heat capacity is zero at $T > T_c$. Moreover, from Eq. (16.26), by evaluating the derivative dM_s/dT, we can see that $C_V \to 0$ at $T \to 0$. At $T = T_c$ the heat capacity exhibits discontinuity. To see that, let us perform calculation in the interval $T < T_c$ using Eq. (16.30),

$$\lim_{T\to T_c} C_V = \frac{3}{2}Nk_B. \tag{16.34}$$

So,

$$\Delta C_V|_{T=T_c} = \frac{3}{2}Nk_B, \tag{16.35}$$

witnessing that the phase transformation at temperature T_c is the second-order phase transition [1]. The Fig. (16.3) schematically illustrates the behavior of heat capacity on temperature in the framework of the molecular field approximation.

Comment: Recall that the heat capacity in the canonical ensemble is related with the fluctuations of internal energy.

[1] Phase transitions according to Ehrenfest propossal have been classified according to the lowest-order derivative of the free energy that becomes discontinuous at the transition. First-order phase transitions are characterized by a discontinuity in the first derivative of the free energy with respect to a thermodynamic variable. Transitions between solid, liquid, and gas phases fall into this category, as they involve a sudden change in density, which is the inverse of the first derivative of free energy with respect to pressure. Second-order phase transitions, in contrast, exhibit continuity in the first derivative of the free energy (e.g., the order parameter, such as magnetization or density), but display a discontinuity in the second derivative, such as heat capacity or compressibility. Detailed description of the theories of phase transition and critical phenomena are given in Ref. [26].

16.2 Molecular field approach to electrolyte solutions

The method developed above has a wider applicability. Now, we would like to illustrate the relevant application in the theory of electrolyte solutions. Consider a multicomponent system of M species with the total number of particles N,

$$\sum_{a=1}^{M} N_a = N. \tag{16.36}$$

The particles of each species a are numbered from 1 to N_a. The particles of each species are charged. The system is considered to be electroneutral because only an electroneutral system of charges can reach equilibrium, otherwise one needs to assume an external the presence of an electric field to equlibrate the system. The condition of electroneutrality is written as,

$$\frac{1}{V}\sum_{a=1}^{M} ez_a N_a = 0, \tag{16.37}$$

where ez_a is the charge of particles belonging to species a and $n_a = N_a/V$ is the average number of particles of species a per unit volume or say the bulk density.

The classical Hamiltonian we would like to deal with depends on momenta (...$\mathbf{p}$...) and coordinates (...$\mathbf{q}$...) of all particles and is as follows,

$$E(...,\mathbf{p}_i,..;...\mathbf{q}_i...) = \sum_{i=1}^{N}\frac{1}{2m}\mathbf{p}_i^2 + \frac{1}{2}\sum_{a,b=1}^{M}\sum_{i=1}^{N_a}\sum_{j=1}^{N_b} ez_a\phi_{ab}(|\mathbf{q}_{ia}-\mathbf{q}_{jb}|), \tag{16.38}$$

where,

$$\phi_{ab}(|\mathbf{q}_{ia}-\mathbf{q}_{jb}|) = \frac{ez_b}{|\mathbf{q}_{ia}-\mathbf{q}_{jb}|}, \tag{16.39}$$

denotes the Coulomb interaction potential between particles i and j belonging to species a and b, respectively. Obvious condition $a,i \neq b,j$ should be satisfied to exclude self interaction.

The essence of the molecular field approximation is to replace the true energy of interaction of any particle with all others by the average value of this energy. In mathematical terms, we would like to apply the approximation,

$$\sum_{b=1}^{M}\sum_{j=1}^{N_b} ez_a\phi_{ab}(|\mathbf{q}_{ia}-\mathbf{q}_{jb}|) \to ez_a < \phi_a >, \tag{16.40}$$

where $< \phi_a >$ formally denotes the average potential of electrostatic field exerted by all particles on the arbitrary (or any) particle of species a at the coordinate $\mathbf{q}_{ia}$. However, the system in question is homogeneous, consequently this average potential solely possesses species index and does not depend on the particles index. With this approximation the average potential energy of Coulomb interactions equals to,

$$< U_N >= \frac{1}{2}\sum_{a=1}^{M}\sum_{i=1}^{N_a} ez_a < \phi_a >= \frac{1}{2}V\sum_{a=1}^{M} ez_a < n_a >< \phi_a > . \tag{16.41}$$

The average electrostatic potential depends on the coordinate, thus upon summation over particles index i in the equation above, we have inserted $< n_a >$ describing an average of the local density of species a of the system at the same coordinate.

In order to obtain necessary potential, one can resort to the Poisson equation of macroscopic electrostatics, see e.g. Appendix 3.11 of Ref. [27],

$$\Delta\phi(\mathbf{q}) = -4\pi\rho(\mathbf{q}), \tag{16.42}$$

where $\rho(\mathbf{q})$ is the function describing the distribution of charges in the system, or say spatial charge density that creates the potential.

We would like to choose the charge distribution in the form,

$$\rho(\mathbf{q}) = \sum_{a=1}^{M} ez_a n_a(\mathbf{q}) \tag{16.43}$$

At this point one should apply Boltzmann statistics. Because the energy of a particle with the charge ez_a at the coordinate q is $ez_a\phi(q)$, the distribution of particles of species a is given by,

$$n_a(q) = n_a \exp\{-ez_a\phi(q)/k_BT\}, \tag{16.44}$$

where n_a is chosen as the bulk density of species a. Thus the electrostatic potential is obtained from the Poisson-Boltzmann equation,

$$\Delta\phi(q) = -4\pi \sum_{a=1}^{M} ez_a n_a \exp\{-ez_a\phi(q)/k_BT\}. \tag{16.45}$$

This is a nonlinear differential equation difficult to solve. However, if the electrostatic potential is weak and/or temperature is high, one can linearize it by expanding the exponent in the r.h.s. into series and keep first two terms of the series,

$$\exp\{-ez_a\phi(q)/k_BT\} = 1 - ez_a\phi(q)/k_BT. \tag{16.46}$$

Recall that the system is electroneutral. Then,

$$\Delta\phi(q) - \frac{4\pi e^2}{k_BT} \sum_{a=1}^{M} n_a z_a^2 \phi(q) = 0. \tag{16.47}$$

Introduce notation,

$$\kappa^2 = \frac{4\pi e^2}{k_BT} \sum_{a=1}^{M} n_a z_a^2; \quad \kappa = 1/r_D, \tag{16.48}$$

where r_D is called as the radius of Debye and κ is termed as the inverse Debye radius.

Restrict to the case when there is a central ion at the origin (it should be an ion belonging to species a with the charge ez_a) and the field then has spherical symmetry w.r.t. the central ion. Then, the potential depends only on the distance r in spherical coordinates. The relevant part of the Laplace operator can be used to transform the equation above into,

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \frac{\partial\phi(r)}{\partial r}) = \kappa^2\phi(r) \tag{16.49}$$

and obtain solution of the form,

$$\phi(r) = A\frac{1}{r}e^{-\kappa r}, \tag{16.50}$$

where A is the constant necessary to determine. The unknown constant is, $A = ez_a$ because of the boundary condition or in other words because we have chosen a given ion at the coordinate center. Recall that Coulomb potential of the particle on its own (or bare, or self-interaction potential), is $\phi_a(r) = ez_a/r$. Thus, the electrostatic potential created by all particles at the coordinate of a given one of our interest is,

$$\phi_a(r) = ez_a\frac{1}{r}e^{-\kappa r}. \tag{16.51}$$

So, what do we have is the electrostatic potential created by the given particle belonging to species a and located at the origin of the coordinate system and by all other charged particles belonging to different species distributed according to Boltzmann statistics. Eq. (16.51) determines the screened potential, rather than the bare Coulomb. The screened potential decays faster than its Coulomb counterpart because each particle creates an "atmosphere" around it. Within this atmosphere, the charges of opposite sign w.r.t. to the central ion charge, prevail. Due to the application of Boltzmann statistics for the charge distribution one can interpret the screened potential as an *average* potential created at a given point.

In order to find the molecular field potential and to apply the approximation defined by Eq. (16.41), we need to rest the self-potential of the central particle from the expression in Eq. (16.51),

$$< \phi_a >= \{ez_a \frac{1}{r} e^{-\kappa r} - ez_a \frac{1}{r}\}. \tag{16.52}$$

Next, in the spirit of linearization of the original Poisson-Boltzmann equation, the screened potential should be expanded into series and the limit $r = 0$ should be performed. Thus,

$$< \phi_a >= -\frac{ez_a}{r_D}. \tag{16.53}$$

We apply then this result in evaluation of the electrostatic contribution into internal energy of the system according to Eq. (16.41). In addition, the average density multiplier is approximated by its bulk value as in Boltzmann distribution given by Eq. (16.44). Then, the result for the internal energy is,

$$U_N^{(el)} = -\frac{V}{2} \frac{1}{4\pi\beta} \kappa^3. \tag{16.54}$$

The approach outlined above corresponds to the Debye-Hückel theory of electrolyte solutions explained in very detail in chapter 15 of Ref. [7] by D. McQuarrie. In this kind of theory only the solute ions are considered explicitly, the effects of solvent may or should be incorporated by renormalization of the charge density distribution by the multiplier $1/\varepsilon$, where ε denotes the dielectric constant of a solvent medium.

17. Adsorption

17.1 Langmuir adsorption

One, pedagogically motivated, application of Boltzmann distribution is the description of Langmuir adsorption.

In general terms we deal with a rather complex system consisting of two subsystems, a fluid and a solid. These two parts of the entire system are put in contact. The simplest model permitting to get insight into the phenomenon of adsorption is the so-called Langmuir model.

According to the model, a solid surface exposed to the adsorption of a fluid is the surface plane containing M identical adsorption sites. The configuration of the sites is not specified. It is assumed that each of the sites can adsorb at most one fluid molecule (obviously the site can remain empty as well). The adsorbed molecules are considered to be independent of one another, i.e. it is an ideal gas of molecules. For simplicity, it is assumed that a molecule, if adsorbed, has energy, $-\varepsilon_0$, relative to its energy in the fluid phase. It can be interpreted as the activation energy. Thus, adsorption of fluid species on the solid surface is energetically preferred comparing to the permanence in the fluid. However, only a monomer layer can by formed according to the assumed setup.

Prior statistical mechanical description of the model, let us resort to the arguments coming from the simplified consideration of kinetics of adsorption. First, introduce the parameter characterizing coverage of the substrate surface by adsorbed molecules, θ, $\theta = N_{ads}/M$, where N_{ads} is the average number of occupied sites. The coverage is a function of time from the point of view of chemical kinetics. In terms of chemical reactions, one can view the process of adsorption as,

$$S + A \rightleftharpoons SA, \tag{17.1}$$

where, the right and left harpoons in terms of rate constants are given by, k_1 (adsorption of A particles or say the formation of SA "species") and k_{-1} (desorption of A particles from the surface sites). At equilibrium, the adsorption, $R_{ads} = k_1 M(1-\theta)P$, and desorption, $R_{des} = k_{-1}M\theta$, rates should satisfy $d\theta(t)/dt = 0$, or in mathematical terms,

$$k_1 M(1-\theta)P = k_{-1}M\theta. \tag{17.2}$$

The state of the adsorbate is characterized by the pressure, P, whereas the solid surface state is described by θ and M, or equivalently by the fraction of "empty" sites, $M(1-\theta)$

and of occupied $M\theta$ sites. Then, the equilibrium surface coverage is,

$$\theta = \frac{KP}{1+KP}, \tag{17.3}$$

where $K = k_1/k_{-1}$ is the equilibrium rate constant. The ratio of rate constants, K, depends on temperature obviously, i.e. $K = K(T)$.

Proceed now to the theoretical considerations. The state of a fluid is determined by the chemical potential and temperature. Such setup is equivalent to the choice of pressure as the external variable, but is more convenient. If one would like to describe adsorption in pores or in disordered porous media, rather than on a single surface. Let the fluid is at thermodynamic conditions given by certain values of μ and T. First, we would like to specify the canonical partition function of N adsorbed molecules. It can be written as,

$$Q_N = \frac{M!}{N!(M-N)!} e^{N\varepsilon_0/k_BT} [q_{ad}^{(in)}(T)]^N. \tag{17.4}$$

The first multiplier $M!/N!(M-N)!$ describes the number of ways the N fluid molecules can be distributed among M identical adsorbing sites. The second multiplier shows that N molecules are in adsorbed state. Finally, the third multiplier is introduced to emphasize that each of adsorbed molecules has internal degrees of freedom yielding the partition function $q_{ad}^{(in)}$.

Having in mind, that imaginary adsorption experiment is performed from a fluid with a given value of chemical potential, and that the number of adsorbed molecules is determined by μ, we are interested in the grand partition function,

$$\Xi = \sum_{N=0}^{M} \frac{M!}{N!(M-N)!} [q_{ad}^{(in)}(T)]^N e^{N(\varepsilon_0+\mu)/k_BT}. \tag{17.5}$$

Explicit expression for the grand partition function follows from the binomial theorem, $((x+y)^M = \sum_{N=0}^{M} \frac{M!}{N!(M-N)!} x^{(M-N)} y^N)$,

$$\Xi = \left\{1 + q_{ad}^{(in)}(T) e^{(\varepsilon_0+\mu)/k_BT}\right\}^M. \tag{17.6}$$

The knowledge of Ξ permits to evaluate the average number of adsorbed molecules straightforwardly,

$$\overline{N} = -\left(\frac{\partial\Omega}{\partial\mu}\right) = k_BT\left(\frac{\partial ln\Xi}{\partial\mu}\right), \tag{17.7}$$

to get,

$$\frac{\overline{N}}{M} = \frac{q_{ad}^{(in)}(T) e^{(\varepsilon_0+\mu)/k_BT}}{1 + q_{ad}^{(in)}(T) e^{(\varepsilon_0+\mu)/k_BT}}, \tag{17.8}$$

which is called as the Langmuir adsorption isotherm. It provides the description of the coverage, $\overline{N}/M$, or say occupancy of the solid adsorbing surface by fluid particles, as a function of external chemical potential describing the state of the bulk fluid phase. The temperature dependence of coverage at a given value of the chemical potential can be studied as well.

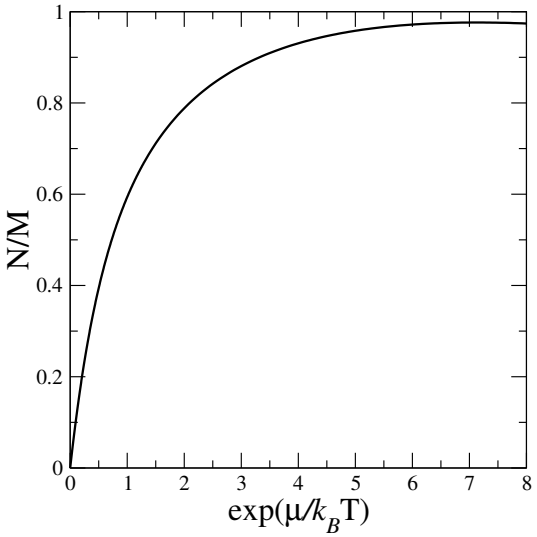


Fig. 17.1: Langmuir adsorption isotherm.

Schematic plot of the dependence of coverage on fluid activity, $\exp(\mu/k_BT)$, is shown in the Fig. (17.1). At a high values of activity, we observe trends for saturation, i.e. the approach to the complete occupancy of the adsorbing sites by fluid species. On the other hand, in the limiting case of small activity, we assume that the denominator of Eq. (17.8) is dominated by the first term, and becomes almost equal to unity (recall that if the system is close to the ideal gas the chemical potential is negative and big by magnitude). Then, Eq. (17.8) can be re-arranged in the following form,

$$\mu = -\varepsilon_0 - k_BT\ln\left(q_{ad}^{(in)}\right) + k_BT\ln\left(\frac{\overline{N}}{M}\right). \tag{17.9}$$

This result can be combined with the Eqs. (13.11) and (13.12) given for the chemical potential for the fluid phase considered as a dilute, almost ideal, gas,

$$\mu = k_BT\ln\left[\frac{\lambda^3}{k_BTq_f^{(in)}}\right] + k_BT\ln(P), \tag{17.10}$$

where $\lambda = h/\sqrt{2\pi mk_BT}$ is the thermal de Broglie length, $q_f^{(in)}$ is the partition function of the internal degrees of freedom of a single particle in the bulk fluid phase and P is the external pressure. Eqs. (17.9) and (17.10), if combined, yield the result $\frac{\overline{N}}{M} = K(T)P$. On the other hand, if the chemical potential from Eq. (17.9) is inserted into Eq. (17.8), on arrives at Langmuir adsorption isotherm,

$$\frac{\overline{N}}{M} = \frac{K(T)P}{1+K(T)P}, \tag{17.11}$$

where,

$$K(T) = \left(\frac{q_{ad}^{(in)}\lambda^3}{q_f^{(in)}k_BT}\right)\exp\left(\frac{\varepsilon_0}{k_BT}\right). \tag{17.12}$$

It can be used to describe adsorption of a dilute gas on a solid substrate. The coverage, $\overline{N}/M$, in such a form is the function of an external pressure describing the state of a gas. However, one can interpret Eq. (17.11) differently. Imagine that the conditions of adsorption experiment are known, i.e. P and T, and that $K(T)$ is experimentally available.

Then, the left had side provides knowledge about the total surface area that the solid makes available for adsorption in terms of $M/\overline{N}$.

The Langmuir adsorption isotherm is very simple, it served for interpretation of adsorption for several years. However, many efforts have been focused in the better description of adsorption phenomena. Along this line of research the so-called BET isotherm has been developed, named after Brunauer, Emmett and Teller. One of the principal differences of this approach comparing to Langmuir, is that the possibility of adsorption of a pile (or column) of fluid particles by a single surface site is assumed. In other words, multilayer adsorption can be captured. To do that, the model assumes a set of energetic states for a particle dependent on the number of adsorbed layer, $-\varepsilon_0, -\varepsilon_1, ...$ Then, if one restricts to the adsorption of a dilute gas, Eq. (17.11) is modified to the BET form,

$$\frac{\overline{N}}{M} = \frac{C(T)\left(\frac{P}{P_s}\right)}{\left(1-\left(\frac{P}{P_s}\right)\right)\left[1+(C(T)-1)\left(\frac{P}{P_s}\right)\right]}, \tag{17.13}$$

where $P_s = P_s(T)$ is the saturated vapor pressure at a temperature T and $C(T)$ is the parameter related to $K(T)$. In the limit of small P/P_s and large $C(T)$, one can rewrite this equation in the Langmuir form,

$$\frac{\overline{N}}{M} = \frac{C(T)\left(\frac{P}{P_s}\right)}{1+\left(C(T)\left(\frac{P}{P_s}\right)\right)} = \frac{K(T)P}{1+K(T)P}, \tag{17.14}$$

if $C(T) = K(T)P_s$. Thus, at low temperature, when P_s is quite low, BET type adsorption isotherm reduces to the form of Langmuir. This is illustrated in the figure below. If the curve given by Eq. (17.13) is plotted schematically it describes multilayer coverage in contrast to Langmuir shape.

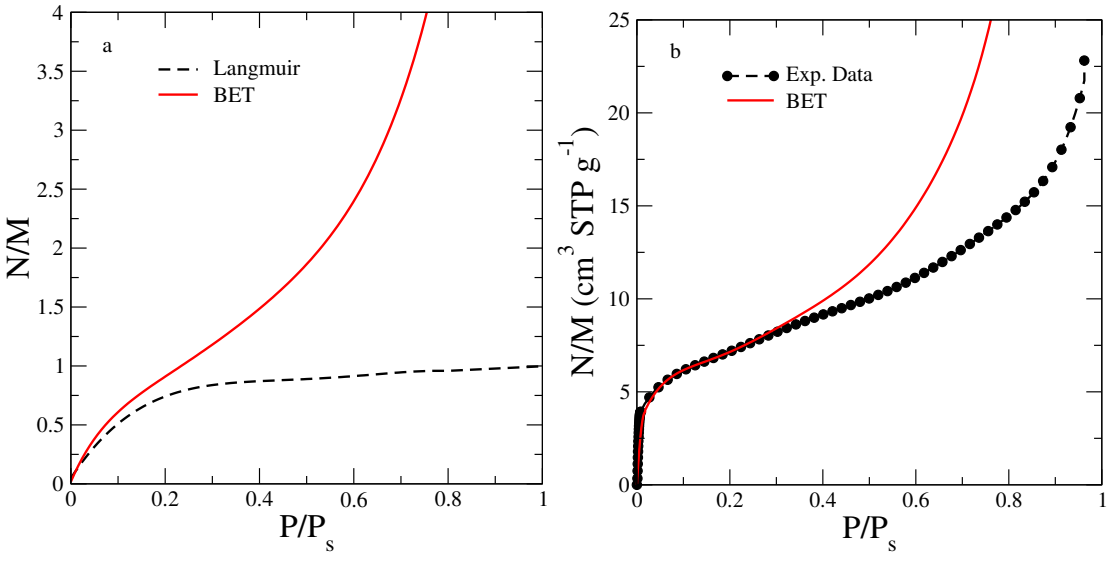


Fig. 17.2: Panel (a): BET versus Langmuir adsorption isotherms. Panel (b): Experimental data and corresponding BET adsorption isotherms. The Panel (b) of the figure is redone from the original in Ref. [28]. The experimental points are from Ref. [29]. The adsorbed amount diverges at $P/P_s \to 1$. Red curve follows from a version of fitting using BET approximation.

To summarize, one can see a conceptual difference of the BET isotherm comparing to Langmuir one, the former requires knowledge of the equation of state of the bulk fluid to

involve the variable P_s. The Brunauer–Emmett–Teller approximation even nowadays is popular to interpret e.g. the nitrogen adsorption isotherms for determining the specific surface area. A qualitatively correct descriptions can be obtained within this method, see Fig. (17.2) for illustration of theoretical, BET type results, and experimental data. Nevertheless, inaccurate description of the growth of macroscopic film precludes accurate determination of the wetting temperature of the solid substrate.

18. Fermi-Dirac and Bose-Einstein statistics

18.1 Ideal gas obeying Fermi-Dirac or Bose-Einstein statistics

If one would like to search for the statistical peculiarities of the distribution of particles over different energetic states, the nature of particles should be taken into account besides thermodynamic conditions of the system. Specifically, we know from the Pauli exclusion principle of quantum mechanics that if two identical particles have half-integer spin angular moment (like electrons, ^{3}He atoms for example) they cannot occupy the same single particle state. In other words, the occupancy of the single particle state is either 0 or 1. This is a situation that leads to Fermi-Dirac statistics, that applies only to a quantum system of non-interacting fermions.

Assume, that the single particle energy states have energy $\varepsilon_1, \varepsilon_2, ...$ and that the number of particles in the state m is n_m. If the system is considered in the canonical ensemble, the total number of particles in the system, N, is fixed. In fact it is the sum over the number of particles in all one-particle states,

$$N = \sum_m n_m. \tag{18.1}$$

Besides, the energy of a single given quantum state of many-particle system is,

$$E = n_1\varepsilon_1 + n_2\varepsilon_2 + ... = \sum_m n_m\varepsilon_m. \tag{18.2}$$

Under such circumstances the canonical partition function is the summation over all possible states of the many-particle system,

$$Q_N = \sum_{\{n_m\}} \exp\{-\beta(n_1\varepsilon_1 + n_2\varepsilon_2 + ...)\}, \tag{18.3}$$

under the condition of Eq. (18.1), i.e. that the sets of numbers in n_m sum to N.

However, the Fermi–Dirac distribution is most easily derived from the grand canonical ensemble. In this ensemble, the system is able to exchange energy and exchange particles with a reservoir, where the temperature and chemical potential are fixed by the reservoir. Due to the absence of interactions between fermions, each available single-particle level forms a separate thermodynamic system in contact with the reservoir. In other words, each single-particle level is a separate, tiny grand canonical ensemble.

The grand partition function of fermions follows from Eq. (18.3) straightforwardly,

$$\Xi = \sum_{N=0}^{\infty} e^{\beta\mu N} Q_N = \sum_{N=0}^{\infty} e^{\beta\mu N} \sum_{\{n_m\}} \exp\{-\beta(n_1\varepsilon_1 + n_2\varepsilon_2 + ...)\}, \tag{18.4}$$

the inner sum now, does not contain any restriction for summation because of the outer sum over different values of the number of particles. Hence, the partition function is,

$$\begin{aligned}\Xi &= \sum_{n_1=0}^{1}\sum_{n_2=0}^{1} \times ... \times \sum_{n_m=0}^{1} \exp\{\beta[n_1(\mu-\varepsilon_1) + n_2(\mu-\varepsilon_2) + ...]\} \\ &= \prod_m (1 + \exp\{\beta(\mu - \varepsilon_m)\}).\end{aligned} \tag{18.5}$$

where $(\partial \ln\Xi/\partial\mu) = -(\partial \ln\Xi/\partial\varepsilon_m)$. With this result we can obtain the probable number of particles (or the average number of particles or occupancy) $\overline{n}_s$ in the energy level ε_s,

$$\overline{n}_s = -\left(\frac{\partial\Omega}{\partial\mu}\right) = k_B T \quad \left(\frac{\partial \ln\Xi}{\partial\mu}\right) = -k_B T\left(\frac{\partial \ln\Xi}{\partial\varepsilon_s}\right) = \frac{1}{e^{\beta(\varepsilon_s-\mu)}+1}. \tag{18.6}$$

This is the Fermi-Dirac occupancy distribution. Schematic graph of the occupancy of levels with different energy at zero and non-zero temperature is shown in the Fig. (18.1) below.

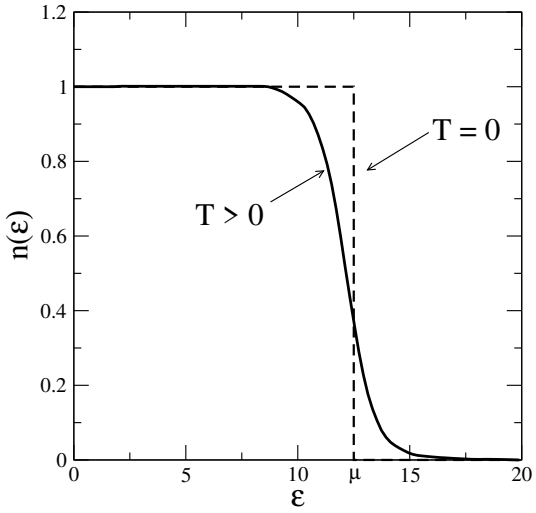


Fig. 18.1: Fermi-Dirac distribution of the occupancy of energetic states.

At zero temperature all energy levels are occupied with the probability equal to 1 up to $\varepsilon = \mu$. On the other hand there are no particles excited to the energy level above μ. At non-zero temperature a fraction of particles occupy the energy levels above μ.

It is worth mentioning that the average for the total number of particles at a given chemical potential can be obtained by summation over levels as follows,

$$\overline{N} = \sum_s \overline{n}_s = \sum_s \frac{1}{e^{\beta(\varepsilon_s-\mu)}+1}. \tag{18.7}$$

These, quite general relations can be used in several applications. More specifically, we would like to illustrate an application to obtain thermodynamics properties of ideal Fermi gas. For a system of N non-interacting fermions, e.g. electrons, the one-particle wave functions are plane waves. A complete set of quantum numbers is given by the components of the wave vector $\mathbf{k}$ (k_x, k_y, k_z) and the projection of spin moment along z-axis, m_z. As we have seen above, the evaluation of the partition function in Eq .(18.5),

requires summation over all energetic states. Such summation over states reduces then to the summation over states with different wave vectors,

$$\sum_r ... \to (2s+1)\sum_{\mathbf{k}} ..., \tag{18.8}$$

the prefactor takes into account possible values of the spin quantum number, recall that the energy is degenerate. If the electrons are in the cubic box with the volume $V = L_x L_y L_z$, the periodic boundary conditions for the wave functions require, $k_x L_x = 2\pi n_x$, cf. Eq. (7.16) of chapter 7 concerned with the simplest problems of quantum mechanics. Therefore, each state in the wave-vector space has an average volume $\Delta k = (2\pi)^3/(L_x L_y L_z)$.

Consequently, if the volume of the system is macroscopic,

$$\sum_r ... \to (2s+1)\frac{1}{\Delta k}\int d^3k = (2s+1)\frac{V}{(2\pi)^3}\int d^3k = (2s+1)\frac{V}{h^3}\int d^3p, \tag{18.9}$$

where the momentum p, equals to $\hbar k$. Thus, the grand thermodynamic potential,

$$\Omega(\mu,T,V) = -k_B T \ln \Xi(\mu,T,V) = -k_B T \sum_i \ln(1+\exp\{\beta(\mu-\varepsilon_i)\}) \tag{18.10}$$

can be written as an integral,

$$-\beta\Omega(\mu,T,V) = (2s+1)\frac{V}{(2\pi)^3}4\pi\int_0^\infty dk k^2 ln[1+z\exp(-\beta\hbar^2k^2/2m)], \tag{18.11}$$

where the activity, z, is $z = \exp(\beta\mu)$ and $\varepsilon_i \to \varepsilon_k = \hbar^2k^2/2m$.

$$x = \hbar k\sqrt{\frac{\beta}{2m}}, \quad k^2dk = \left(\frac{2m}{\beta\hbar^2}\right)^{3/2} x^2dx, \tag{18.12}$$

to obtain,

$$-\beta\Omega(\mu,T,V) = (2s+1)\frac{V}{2\pi^2}\left(\frac{2m}{\beta\hbar^2}\right)^{3/2}\int_0^\infty dx x^2 \ln[1+z\exp(-x^2)]. \tag{18.13}$$

Further simplification can be reached by using the de Broglie thermal wavelength, $\lambda = \sqrt{2\pi\beta\hbar^2/m}$,

$$-\beta\Omega(\mu,T,V) = (2s+1)\frac{4V}{\lambda^3\pi^{1/2}}\int_0^\infty dx x^2 \ln[1+z\exp(-x^2)]. \tag{18.14}$$

Now, the principal issue is the evaluation of the integral in the equation above. To do that, one should use Taylor expansion of the logarithm. Then the desired integration can be performed as follows,

$$\begin{aligned}\int_0^\infty dx x^2 \ln[1+z\exp(-x^2)] &= \sum_{n=1}^\infty (-1)^{(n+1)}\frac{z^n}{n}\int_0^\infty dx x^2 e^{-nx^2} \\ = \sum_{n=1}^\infty (-1)^{(n+1)}\frac{z^n}{n}(-\frac{d}{dn}\int_0^\infty dx e^{-nx^2}) &= \sum_{n=1}^\infty (-1)^{(n+1)}\frac{z^n}{n}\frac{1}{2}(-\frac{d}{dn}\sqrt{\frac{\pi}{n}}) \\ &= \frac{\sqrt{\pi}}{4}\sum_{n=1}^\infty (-1)^{(n+1)}\frac{z^n}{n^{5/2}}.\end{aligned} \tag{18.15}$$

The sum in the last line is related to the Γ function,

$$f_{5/2}(z) = \sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{z^n}{n^{5/2}} = \frac{4}{\sqrt{\pi}} \int_0^{\infty} dx x^2 \ln[1 + z\exp(-x^2)]. \tag{18.16}$$

Then, finally the grand thermodynamic potential of an ideal Fermi gas as a function of activity, is,

$$\beta\Omega(\mu, T, V) = -\frac{2s+1}{\lambda^3} V f_{5/2}(z). \tag{18.17}$$

We have derived previously the thermodynamic relation $\Omega = -PV$. Consequently, the $P = P(\mu, T)$ projection of the equation of state is,

$$\frac{P}{k_B T} = \frac{2s+1}{\lambda^3} f_{5/2}(z). \tag{18.18}$$

The average number of particles has been defined above in Eq. (18.7). However, that formal expression can be elaborated by using the result for Ω. Namely,

$$\overline{N} = -\left(\frac{\partial\Omega}{\partial\mu}\right) = z\left(\frac{\partial}{\partial z}\ln\Xi\right). \tag{18.19}$$

Then,

$$n = \frac{\overline{N}}{V} = \frac{2s+1}{\lambda^3(T)} f_{3/2}(z), \tag{18.20}$$

where,

$$f_{3/2}(z) = z\frac{d}{dz} f_{5/2}(z). \tag{18.21}$$

This Eq. (18.20) represents the $z = z(n, T)$ projection of the equation of state for ideal Fermi gas. Other thermodynamic properties can be obtained from the expression of the grand thermodynamic potential as well. Several important problems, like the heat capacity of electron gas, can be explored, in addition to the brief account of this chapter above.

The other case of quantum statistics refers to the particles that have integer total spin angular momentum each (for example photons, neutrons, etc.). For this kind of particles any number of them can occupy the same single particle quantum state. Again we refer to a system of non-interacting particles of this kind. Then, the partition function in the grand canonical ensemble can be found as we have done above for fermions, but without restriction for occupation of each level,

$$\Xi = \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \times ... \times \exp\{\beta[n_1(\mu-\varepsilon_1) + n_2(\mu-\varepsilon_2) + ...]\} = \prod_s \frac{1}{(1-\exp\{\beta(\mu-\varepsilon_s)\})}, \tag{18.22}$$

because each sum in the left hand side represents a geometric series of the relevant variable. The average number of particles occupying certain state, s, then is,

$$\overline{n}_s = k_B T\left(\frac{\partial ln\Xi}{\partial\mu}\right) = \frac{1}{e^{\beta(\varepsilon_s-\mu)} - 1}. \tag{18.23}$$

This expression is known as Bose-Einstein statistics. The relation between the chemical potential and the average number of particles in the entire system, then follows straightforwardly,

$$\overline{N} = \sum_s \frac{1}{e^{\beta(\varepsilon_s-\mu)} - 1}. \tag{18.24}$$

We do not elaborate the applications of Bose-Einstein statistics for the moment, however.

Recall now our discussion concerning a classical system where the energy levels are distributed continuously and the average occupation of a level is much less than unity, $n_s << 1$. To get this limiting case from both the Fermi-Dirac and Bose-Einstein statistics, the following inequality should hold,

$$e^{\beta(\varepsilon_s - \mu)} >> 1. \tag{18.25}$$

Then, the average occupancy of a state is,

$$\overline{n}_s = e^{\beta(\mu - \varepsilon_s)}, \tag{18.26}$$

and represents Maxwell-Boltzmann statistics.

To summarize, we have derived rules of Fermi-Dirac (FD) and Bose-Einstein (BE) statistics. The former, FD statistics, has been applied to derive some thermodynamics properties of an ideal gas of fermions. On the other hand, the application of the BE statistics is given in the following chapter.

19. Blackbody radiation

19.1 Ideal photon gas

To begin with, let us comment briefly the system we are dealing with. A photon gas is a gas-like collection of photons, which has many of the same properties of a conventional gas including pressure, temperature, and entropy. The most common example of a photon gas at equilibrium is black-body radiation.

Photons are particles known as bosons, i.e. the particles that obey Bose-Einstein statistics and that have integer spin momentum. In general, a single-component gas of bosons is uniquely described by three variables such as the temperature, volume, and the number of particles. However, for a black body radiation the energy distribution is established by the interaction of photons with matter, usually with the walls of the container. Within this setup, the number of photons is not conserved.

In a photon gas, there will exist an equilibrium distribution, but photons do not collide with each other (except under extreme conditions), so the equilibrium distribution must be established by other means. The most common way that an equilibrium distribution is established is by the interaction of the photons with matter. A photon may collide with an electron in the wall, exciting it to a higher energy state, removing a photon from the photon gas. This electron may drop back to its lower level in a series of steps, each one of which releases an individual photon back into the photon gas. If the photons are absorbed and emitted by the walls of the system containing the photon gas, and the walls are at a particular temperature, then the equilibrium distribution for the photons will be a black-body distribution at that temperature.

Now proceed with thermodynamic arguments. It is known from electrodynamics, that the equation of state for electromagnetic radiation is,

$$E = 3PV, \tag{19.1}$$

where $P = |\vec{P}|$, and E is the energy of radiation. This equation does not contain the temperature variable explicitly. Thus it is of interest if and how the internal energy density $E/V = 3P$ depends on temperature. As we know already, the pressure follows from thermodynamic relation, $P = -(\partial A/\partial V)_T$. For common ideal gas $P(N,V,T) = (N/V)k_BT$. However, for electromagnetic radiation, the number of particles is not fixed, so $P = P(V,T)$. The pressure should be an intensive variable, thus it cannot depend on V, and consequently $P = P(T)$. Therefore, using integration of just mentioned Maxwell relation for pressure,

we obtain the Helmholtz free energy as,

$$A = -PV = -\frac{1}{3}E. \tag{19.2}$$

Now, we see that for classical electromagnetic radiation, the free energy is negative and besides, it is proportional to the total energy. Since, $A = E - TS$ and entropy is $S = -(\partial A/\partial T)_V$, we obtain the first-order differential equation,

$$A = -3A + T\left(\frac{\partial A}{\partial T}\right)_V. \tag{19.3}$$

The solution of the equation above is,

$$A \propto T^4. \tag{19.4}$$

Consequently, $E \propto T^4$ as well. Since the energy is extensive, $E = 3P(T)V = const \times VT^4$. Conventionally, this dependence is written as,

$$\frac{E}{V} = 4\sigma \times \frac{1}{c}T^4, \tag{19.5}$$

where c is the velocity of light. The dependence of E on T^4 is known as Stefan's law. It is important to note, that thermodynamics alone is unable to determine the proportionality constant σ.

Final observation using thermodynamic relations is the following. The Gibbs free energy, G, is,

$$G = A + PV = -PV + PV = 0. \tag{19.6}$$

Since $G = \mu N$, we conclude that the chemical potential of electromagnetic radiation is zero, $\mu = 0$. Interpretation of this fact can be given in the spirit of mass action law. Since radiation can be produced by vibrating atoms, the "reaction" atom $\rightarrow$ atom + radiation is allowed. However, the sum of chemical potentials on the two sides of a reaction should be equal, therefore we conclude that the chemical potential for electromagnetic radiation must vanish.

Let proceed to the statistical mechanical treatment now. To begin with, let us assume that the energy states $\varepsilon_1, \varepsilon_2, ...$ can be occupied by $n_1, n_2, ...$ particles, where n_i is any positive integer, $0, 1, 2, ...$ (the state can be unoccupied as well). Recall, that the total number of particles is not fixed.

For this case, the grand canonical partition function of the system is,

$$Q = \sum_{n_1,n_2,...} e^{-\beta n_1\varepsilon_1 - \beta n_2\varepsilon_2 ...} = \prod_m (\sum_{n_m=0}^{\infty} e^{-\beta n_m \varepsilon_m}) = \prod_m \frac{1}{(1 - e^{-\beta\varepsilon_m})}. \tag{19.7}$$

where the summation is performed over all energy states with arbitrary number of particles in each state.

It makes possible to calculate the average energy as we have done in several occasions in previous chapters,

$$< E >= -\frac{\partial ln Q}{\partial \beta} = \sum_n \varepsilon_n \frac{1}{e^{\beta\varepsilon_n} - 1}. \tag{19.8}$$

This expression matches the general formula, $< E >= \sum \varepsilon_i < n_i >$, with $< n_i >= (e^{\beta\varepsilon_i} - 1)^{-1}$ according to Bose-Einstein statistics at $\mu = 0$. On the other hand, one can get the pressure from,

$$P = \frac{1}{\beta}\left(\frac{\partial ln Q}{\partial V}\right)\Big|_T. \tag{19.9}$$

To do that, consider a cubic volume, $V = L^3$, and besides recall the relations involving photons energy, momentum and wave vector, $\mathbf{p} = \hbar\mathbf{k}$, $\varepsilon = \hbar\omega$, $|\mathbf{p}| = \hbar\omega/c$ (discretize the wave vector as we did in chapter 8 for a particle in a cubic box). Then, $\varepsilon_n = \pi\hbar cn/L$. Now we are able to calculate pressure as follows (one can easily do that on his own),

$$P = \frac{1}{\beta}\left(\frac{\partial lnQ}{\partial \varepsilon_n}\right)\left(\frac{\partial \varepsilon_n}{\partial L}\right)\left(\frac{\partial L}{\partial V}\right) = \frac{1}{3V} < E > . \tag{19.10}$$

This result, $< E >= 3PV$, is identical to what was obtained solely from thermodynamic arguments slightly above.

From the expression for the partition function, we can extract one more result. Namely, because,

$$\frac{\partial lnQ}{\partial \varepsilon_s} = -\beta\frac{1}{Q}\sum_{n_1,n_2,\ldots} n_s e^{-\beta n_1\varepsilon_1 - \beta\varepsilon_2 \ldots} \tag{19.11}$$

it follows that the average occupation number of the energy level with energy ε_s is,

$$\overline{n}_s = \frac{e^{-\beta\varepsilon_s}}{1 - e^{-\beta\varepsilon_s}} = \frac{1}{e^{\beta\varepsilon_s} - 1}. \tag{19.12}$$

The result is the so-called Planck distribution. Actually, it is the same result as we have obtained for the BE statistics in the previous chapter, but for the particular case when the chemical potential equals to zero.

Now, we would like to take advantage of the Planck distribution for the occupation number of different states in order to get insight into the energy distribution of particles in the system. To do that resort to the procedure described by Eqs. (18.8) and (18.9) of our previous chapter 18. The number of quantum states within the element of momentum space d^3p for the system with volume V is d^3pV/h^3 (cf. Eq. (18.9)).

Then, the probable number of photons in the element of the wave vector space d^3k per unit volume is,

$$\overline{n}(\varepsilon)(\frac{d^3pV}{h^3})\frac{1}{V} = f(\mathbf{k})d^3k = \frac{1}{(e^{\beta\hbar\omega} - 1)}\frac{d^3k}{(2\pi)^3}. \tag{19.13}$$

In essence, we prepared elements for the integration over relevant variables, rather than the summation over energy states assuming that the system is macroscopic and energy spectrum is continuous.

Equivalent, but more convenient for purposes of our study, equation for the energy distribution function on the frequency of radiation (denote it as $e(\omega)$), follows for the result just above. First, it is necessary to mention that we are looking for the distribution irrespective to the direction of the wave vector, i.e. $d\mathbf{k} = 4\pi k^2 dk = 4\pi\omega^2 d\omega/c^3$, because of $\omega = ck$. The photon energy is $\hbar\omega$, but one needs to take into account two directions of the electric polarization of the wave, thus the energy distribution involves $2\hbar\omega$ rather than $\hbar\omega$. With these preliminaries, the desired expression for the $\varepsilon(\omega)$ can be obtained as follows,

$$e(\omega)d\omega = \frac{4\pi\omega^2 d\omega/c^3}{(e^{\beta\hbar\omega} - 1)}(2\hbar\omega) = \frac{\hbar}{\pi^2c^3}\frac{\omega^3 d\omega}{(e^{\beta\hbar\omega} - 1)}. \tag{19.14}$$

Hence, the l.h.s. $e(\omega)d\omega$ is the energy density (per unit volume) resulting from photons of a gas that have frequency in the interval, $\omega, \omega + d\omega$. It is the probabilistic function because the Planck distribution is involved.

The r.h.s. of Eq. (19.14) has transparent meaning, it can be written as,

$$\frac{\hbar}{\pi^2c^3}\frac{\omega^3 d\omega}{(e^{\beta\hbar\omega} - 1)} = \frac{1}{V}(\hbar\omega)\overline{n}(\omega)g(\omega)d\omega, \tag{19.15}$$

where,

$$g(\omega) = \frac{V}{\pi^2 c^3}\omega^2, \tag{19.16}$$

is the density of states. It provides the measure for integrating over energies. The number of states between ω and $\omega + d\omega$ is $g(\omega)d\omega$.

The plot of the distribution of energy of radiation is shown in Fig. 19.1. It shows the function $e(\omega)$ versus $\hbar\omega/k_BT$, denominated as the blackbody spectral distribution function.

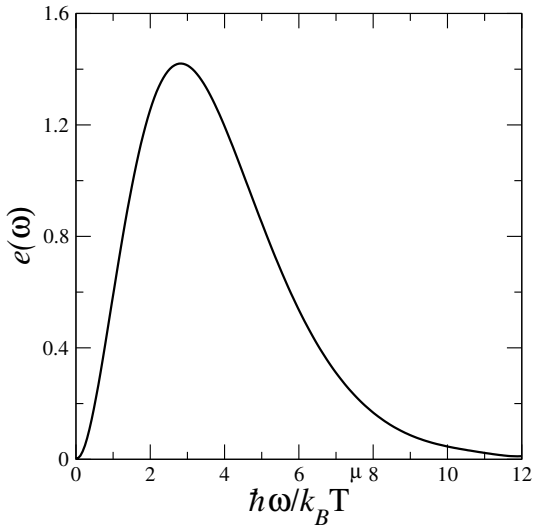


Fig. 19.1: Scaled blackbody spectral distribution function $e(\omega)$ on photon energy $\hbar\omega$ in k_BT units.

One observes that the curve has a maximum at certain frequency, it can be found straightforwardly from $\partial e(\omega)/\partial\omega$. The maximum of radiation energy occurs at,

$$\frac{\hbar\omega_{max}}{k_BT} = 2.821439. \tag{19.17}$$

This relation is known as Wein displacement law for the black body radiation. The concluding remark at this point is that it is possible to determine the temperature of a blackbody by observing the frequency of the maximum in the radiated intensity. In other words, one has a method for the measurement of the temperature of a star or, for example, of the cosmic background radiation.

The radiation frequency can be transformed into the wavelength, $\omega = 2\pi c/\lambda$. For this purpose, introduce the spectral distribution function of the wavelength,

$$e(\lambda)d\lambda = e(\omega)\left|\frac{d\omega}{d\lambda}\right|d\lambda = \frac{8\pi hc}{\lambda^5}\frac{1}{e^{hc/\lambda k_BT} - 1}d\lambda. \tag{19.18}$$

The plot of this kind of distribution function on wavelength is given in Fig. 19.2.

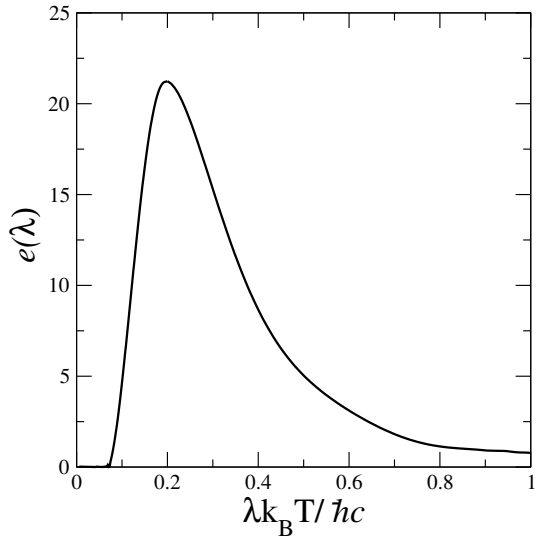


Fig. 19.2: Scaled blackbody spectral distribution function $e(\lambda)$ on photon wavelength.

The wavelength of most intense radiation follows from Eq. (19.17) straightforwardly,

$$\lambda_{max} = \hbar c/(2.821439 k_B T). \tag{19.19}$$

Obviously, the wavelength of most intense radiation is determined by the temperature.

To appreciate this result, let us mention that a black body source will irradiate ultraviolet light (with the wavelength $\approx 10^4$ Å) at $T = 14400K$, visible light (with $\approx 3 \times 10^3$Å) at $T = 4800K$ and finally the radiowaves (with the wavelength $30m$) at $T = 4.8 \times 10^{-5}K$. Illustration of these trends of behavior is given in Fig. 19.3.

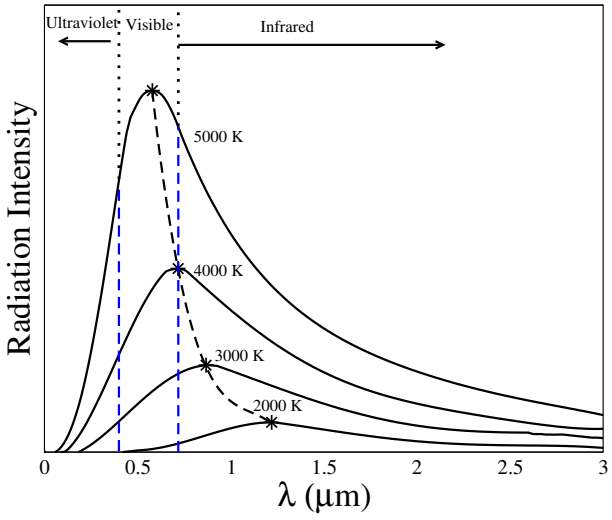


Fig. 19.3: Radiation intensity of a blackbody on wavelength at different temperatures.

It is of interest to obtain the total energy density at a given temperature, by integration over possible frequencies,

$$e_{tot}(T) = \int_0^\infty e(\omega)d\omega = \frac{\hbar}{\pi^2 c^3}\left(\frac{k_B T}{\hbar}\right)^4 \int_0^\infty \frac{x^3}{e^x - 1}dx = \frac{\pi^2}{15}\frac{(k_B T)^4}{(c\hbar)^3}. \tag{19.20}$$

Thus, $e_{tot}(T) \propto T^4$. This dependence is called the Stefan-Boltzmann law. It is worth noting that the last equality in Eq. (19.20) follows from the representation of the integral in the

form of Riemann zeta function. Now, we identify our result from Eq. (19.20) with Eq. (19.5), $e_{tot}(T) = E/V$. Moreover, we see that the quantum statistical mechanics has provided the proportionality factor, σ, that thermodynamics alone is unable to provide,

$$\sigma = \frac{\pi^2}{60}\frac{k_B^4}{c^2\hbar^3}. \tag{19.21}$$

It is known as Stefan's constant.

The heat capacity at constant volume follows straightforwardly for the result for energy density,

$$C_V = \left(\frac{\partial < E >}{\partial T}\right)\Bigg|_V = 16\sigma T^3 V/c. \tag{19.22}$$

It tends to zero with $T \to 0$, in agreement with the third law of thermodynamics. The entropy can be evaluated as well,

$$S = \int_0^T dT' \frac{C_V(T')}{T'} = \frac{16}{3}\sigma T^3 V/c. \tag{19.23}$$

Finally, the total number of photons at a given temperature follows from,

$$N = \int d\omega g(\omega)\frac{1}{(e^{\beta\hbar\omega} - 1)} = 16\pi V\left(\frac{k_B T}{hc}\right)^3 \zeta_3, \tag{19.24}$$

where $\zeta_3 = \zeta(3) = 1.202$. Interestingly, the entropy and N are proportional to each other,

$$S = 3.61 k_B N. \tag{19.25}$$

Thus, each photon has $3.61k_B$ of entropy. Note, that this entropy per photon is independent of temperature. All these thermodynamic properties permit to discuss various aspects of the properties of the cosmic microwave background [30], in comparison to the properties of matter. In this text, however, we are mainly interested in the application of concepts of statistical thermodynamics.

20. Tonks-Takahashi fluids

20.1 Tonks-Takahashi model fluid

In this chapter we consider one example of the calculation of the partition function by using isothermal-isobaric ensemble. This application paves the way to further developments tightly related to the theory of liquids that will become the subject of the following chapters.

Let us recall that the partition function of the isobaric ensemble can be written as (cf. Eq. (12.9) of chapter 12),

$$\Delta(P,T) = \frac{1}{v_0}\int_0^\infty dV(\sum_i e^{-\beta E_i(N,V)})e^{-\beta PV}, \tag{20.1}$$

that transforms into commonly used form, if the summation over energy levels is substituted by integration over momenta and coordinates, according to classical statistical mechanics,

$$\Delta(P,T) = \frac{1}{v_0}\int_0^\infty e^{-\beta PV}Q_N(V,T)dV, \tag{20.2}$$

where v_0 is the small element of volume, and Q_N is the partition function of the canonical ensemble. Therefore, as commonly said, the partition function of isobaric ensemble is the Laplace transform of the canonical partition function. Then the key issue is in the evaluation of the latter explicitly, to perform next the desired transform.

Our attention here is in the classical description of the model fluid. Moreover, for the sake of simplicity, the one-dimensional fluid model is under consideration. Therefore we start from,

$$\Delta = \frac{1}{l_0}\int_0^\infty dLe^{-\beta PL}Q_N(L,T), \tag{20.3}$$

Here, L is the one-dimensional counterpart of the volume of the system.

The canonical partition function can be written as follows,

$$Q_N(L,T) = \frac{1}{N!}\left(\frac{2\pi mk_BT}{h^2}\right)^{N/2}[q_{int}(T)]^N Z_N(L,T), \tag{20.4}$$

where, $q_{int}(T)$ is the partition function of internal degrees of freedom of a molecule except of its kinetic energy describing center of mass motion. Most important is, however, the form of the configurational partition function, Z_N (cf. Eq. (12.9) in chapter 12),

$$Z_N(L,T) = \int_0^L ... \int_0^L dx_1 dx_2 ... dx_N e^{-\beta U_N(x_1,x_2,...,x_N)}, \tag{20.5}$$

where U_N is the potential energy of the inter-particle interactions, and x_i is the coordinate of a particle i.

Now, we would like to specify the inter-particle interactions. It is assumed that the potential energy is the sum of pairwise short-range interactions of the form,

$$u(x_{ij}) = \begin{cases} \infty, & |x_{ij}| < a \\ \psi(x_{ij}), & a < |x_{ij}| < \lambda a. \\ 0, & |x_{ij}| > \lambda a \end{cases} \tag{20.6}$$

Here, $\psi(x)$ is rather arbitrary function (different choices are discussed slightly below), λ is a number between 1 and 2. In this one-dimensional fluid a given particle interacts only with its nearest neighbors. The next nearest neighbor is situated further away than $2a$, i.e. at a distance larger than the range of the potential. Thus, the potential energy of the system of N particles consists of $N-1$ terms and reads,

$$U_N(x_1,x_2,..,x_N) = u(x_{12}) + u(x_{23}) + ... + u(x_{N-1,N}), \tag{20.7}$$

where the abbreviation ij is used to denote $x_{ij} = x_i - x_j$.

However, according to the setup given in the Fig. 20.1, we have two more additional terms, $u(x_{01})$ and $u(x_{N,N+1})$. The box length L is $L = x_{N+1} - x_0 - a$. The volume limits are established by x_0 and $x_{N+1} = L + x_0 + a$.

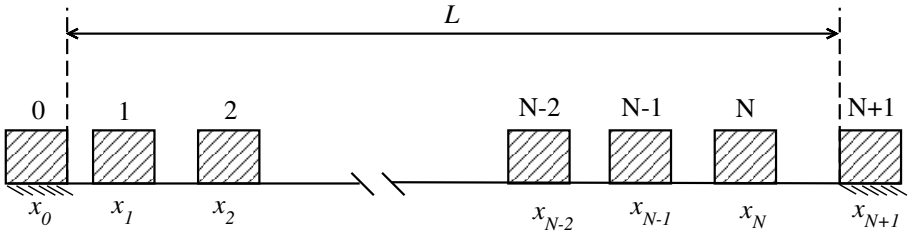


Fig. 20.1: Tonks-Takahashi one-dimensional fluid of N particles confined to an interval $L = x_{N+1} - x_0 - a$. The limits of the system are determined by the same kind of particles fixed at the coordinate x_0 and x_{N+1}, respectively [31].

With this kind of consideration, the auxiliary configuration partion function to the one given by Eq. (20.4), can be written as follows,

$$Z_N(x_0,x_{N+1}) = N! \int_{x_0}^{x_{N+1}} dx_N \int_{x_0}^{x_N} dx_{N-1} ... \int_{x_0}^{x_3} dx_2 \int_{x_0}^{x_2} dx_1 e^{-\beta[u(x_{12})+u(x_{23})+...+u(x_{N-1,N})]} \times e^{-\beta[u(x_{10})+u(x_{N+1,N})]}. \tag{20.8}$$

where $x_0 \le x_1 \le x_2 ... \le x_N \le x_{N+1}$. This kind of ordering of independent integration can be performed because of the assumption of interactions between nearest neighbors only. In spite of simplifications discussed above, explicit evaluation of $Z_N(L,T)$ as a function of L is not straightforward.

Let us attempt it. To perform integrations, introduce first new variables, $y_j = x_j - x_0 - ja$ ($j = 1,2,...,N+1$). Then the auxiliary partition function of the Tonks-Takahashi gas becomes as follows,

$$Z_N(y_{N+1}) = N! \int_0^{y_{N+1}} dy_N g_1(y_{N+1,N}) \int_0^{y_N} dy_{N-1} k(y_{N,N-1}) ... \int_0^{y_3} dy_2 k(y_{32}) \int_0^{y_2} dy_1 k(y_{21}) g(y_1), \tag{20.9}$$

where we introduced notation for the Boltzmann factor, $k(y_{ij}) = \exp[-\beta u(y_{ij})]$, $y_{ij} = y_i - y_j$. The only Boltzmann factors distinguished within this notation, is for the particles, 0 and $N+1$, providing boundaries for the system, $g_1(y_{N+1,N}) = \exp[-\beta u(y_{N+1,N})]$ and $g(y_1) = \exp(-\beta u(y_{1,0}))$ These functions are distinguished, because later we will put the corresponding potentials equal to zero. It is obvious that Eq. (20.8) has the form of an iterated convolution integral. Therefore the Laplace transform method can be used (the Laplace transform of each convolution integral is the product of Laplace transforms of the functions involved in the convolution, see the supplement to this chapter below). The Laplace transform of the function given by Eq. (20.8) is,

$$\int_0^{\infty} dy_{N+1} e^{-s y_{N+1}} Z_N(y_{N+1}) = N!\, g_1(s) [k(s)]^{N-1} g(s). \tag{20.10}$$

Now, we would like to adjust this result to the calculation of the partition function of the isobaric ensemble according to Eq. (20.3). It requires integration of the form,

$$\int_0^{\infty} dL e^{-\beta P L} Z_N(L,T). \tag{20.11}$$

Therefore, we replace y_{N+1} by L in Eq. (20.9). However, the value for L is assumed such, that the $N+1$ particle is excluded from the definition of the volume of the system, or in other words the interaction $u_{N+1,N}$ is not present. Hence,

$$Z_N(L,T) = N!\, [k(s)]^{N-1} g(s). \tag{20.12}$$

On the other hand, the interaction $u_{10} = u(y_{1,0})$ should not be present as well, because the particle with the number 0 is fictitious, it serves only to define the beginning of the coordinate axis. Thus, in the gas we deal with, $u_{10} = u(y_{1,0}) = 0$, and besides the coordinate y_1 (recall it is defined as $y_1 = x_1 - x_0 - a$) should be positive. Consequently, we take $g(y)$ in the form of the Heaviside function,

$$g(y) = \begin{cases} 0, & y < 0, \\ 1, & y \geq 1, \end{cases} \tag{20.13}$$

that has the Laplace transform, $\tilde{g}(s) = e^{cs}/s$ with $c = 0$. The Laplace variable is βP according to Eq. (20.10). Then, the desired partition function is,

$$\Delta(P,T) = \frac{1}{l_0} \left(\frac{2\pi m k_B T}{h^2} \right)^{N/2} [q_{int}(T)]^N \frac{1}{\beta P} [\tilde{k}(\beta P)]^{N-1}, \tag{20.14}$$

where the specific form of the Laplace transform,

$$\tilde{k}(\beta P) = \int_0^{\infty} dt e^{-\beta P t - \beta u(t)}, \tag{20.15}$$

depends on the model of inter-particle interactions $u(t)$.

However, first consider more general relations. The chemical potential of the fluid in question follows from the thermodynamic relation,

$$\mu = -k_B T \left(\frac{\partial ln\Delta}{\partial N} \right)_{P,T}. \tag{20.16}$$

Then,

$$\mu = \mu_{id} - k_B T \ln(\tilde{k}(\beta P)), \tag{20.17}$$

with the ideal contribution,

$$\mu_{id} = -k_B T \ln \left[\left(\frac{2\pi m k_B T}{h^2} \right)^{1/2} q_{int}(T) \right]. \tag{20.18}$$

The equation of state of the one-dimensional fluid involving density of particles can be obtained from the Gibbs-Duhem equation,

$$n^{-1} = \left(\frac{L}{N} \right) = \left(\frac{\partial \mu}{\partial P} \right)_T. \tag{20.19}$$

Now, consider a few specific cases. The simplest one is the hard-rod fluid in which the particles do not interpenetrate each other and possible interparticle attraction is absent, i.e. $\psi(x) = 0$ the equation determining the interaction potential. Then, $\tilde{k}(\beta P) = \exp(-\beta P a)/\beta P$, and inserting this result into Eq. (20.17), one obtains the chemical potential as follows,

$$\mu = \mu_{id} + k_B T \ln(\beta P) + Pa. \tag{20.20}$$

It is the $\mu - P$ projection of the equation of state for hard-rod fluid. The $P - n$ projection of the equation of state follows from Eq. (20.18). It yields,

$$P = \frac{n k_B T}{1 - na} = \frac{N k_B T}{L - Na}. \tag{20.21}$$

This result differs from the EOS of one-dimensional ideal gas, $P = Nk_BT/L$, due to accounting for the particles' dimension or say excluded volume effects.

Two, more sophisticated examples include the effects of interparticle attraction. Namely, the one-dimensional square-well model fluid is determined by $\phi(x) = -\varepsilon$. The Laplace transform of the interaction potential in this case is,

$$\tilde{k}(\beta P) = \frac{1}{\beta P} \left[e^{-\beta P a + \beta \varepsilon} + (1 - e^{\beta \varepsilon}) e^{-\beta \lambda a} \right). \tag{20.22}$$

This model has the following projections of the EOS,

$$\mu = \mu_{id} + k_B T \ln(\beta P a) + Pa - \varepsilon \ln \left[e^{\beta \varepsilon} + (1 - e^{\beta \varepsilon}) e^{-\beta P (\lambda - 1) a} \right], \tag{20.23}$$

and

$$1/n = a + \frac{1}{\beta P} + \frac{(\lambda - 1) a (1 - e^{\beta \varepsilon}) e^{-\beta P (\lambda - 1) a}}{e^{\beta \varepsilon} + (1 - e^{\beta \varepsilon}) e^{-\beta P (\lambda - 1) a}}. \tag{20.24}$$

Finally, one can obtain similar set of results for the model with triangular attraction,

$$\psi(x) = -\varepsilon \frac{(\lambda a - x)}{a(\lambda - 1)}. \tag{20.25}$$

The Laplace transform necessary to calculate thermodynamic properties then is,

$$\tilde{k}(\beta P) = \frac{e^{-\beta P \lambda a}}{\beta P} + \frac{e^{-\beta P a} - e^{-\beta P \lambda a}}{\beta P + \beta\varepsilon/((\lambda - 1)a)}. \tag{20.26}$$

We do not elaborate this model in further details for brevity and because qualitatively new features are not observed comparing to square well fluid.

These two types of projections of the EOS for square-well and triangle-well pair potentials are shown in Fig. 20.2. The density - pressure curve is monotonic for all three models. This means that in spite of the presence of inter-particle attraction in two latter models, no liquid-vapor transition (this would result in non-monotonous behavior) occurs. This feature illustrates what is known as van Hove theorem: a one dimensional fluid cannot undergo a phase transition, if the interparticle interaction is of finite range [32]! (Recall our previous discussions in subsection 15.3 concerning one-dimensional Ising model and implications for the possibility of phase transitions [10]). However, this statement does not discard that anomalous behavior of some thermodynamic properties (recall the density anomaly and other anomalies of real water and within various models) can be observed in one-dimensional models.

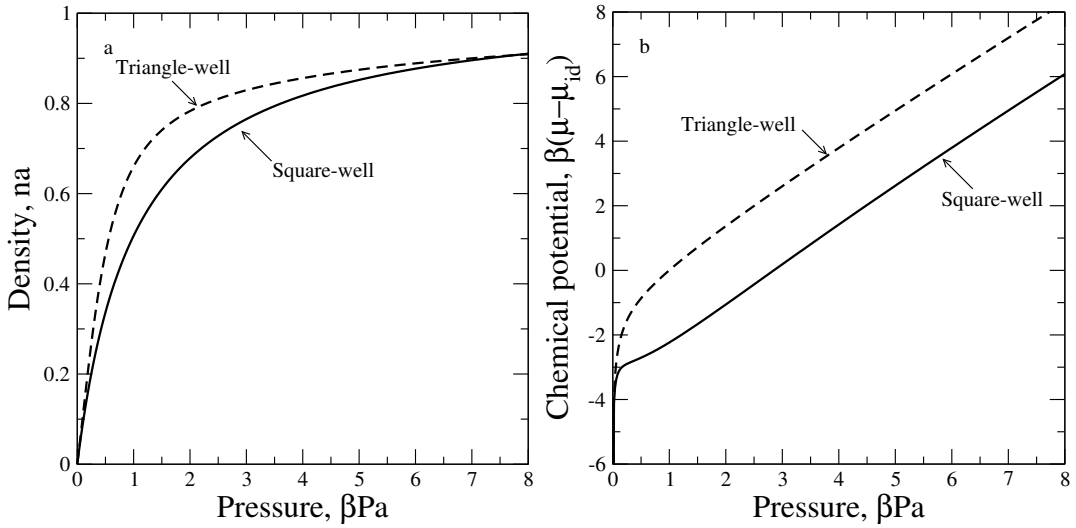


Fig. 20.2: Panel a: Density - pressure projection of the EOS for one-dimensional models of this chapter. Panel b: Chemical potential - pressure projection of the EOS for one-dimensional models of this chapter. For the square-well and triangle-well pair potentials $\beta\varepsilon = 2$, and $\lambda = 2.0$.

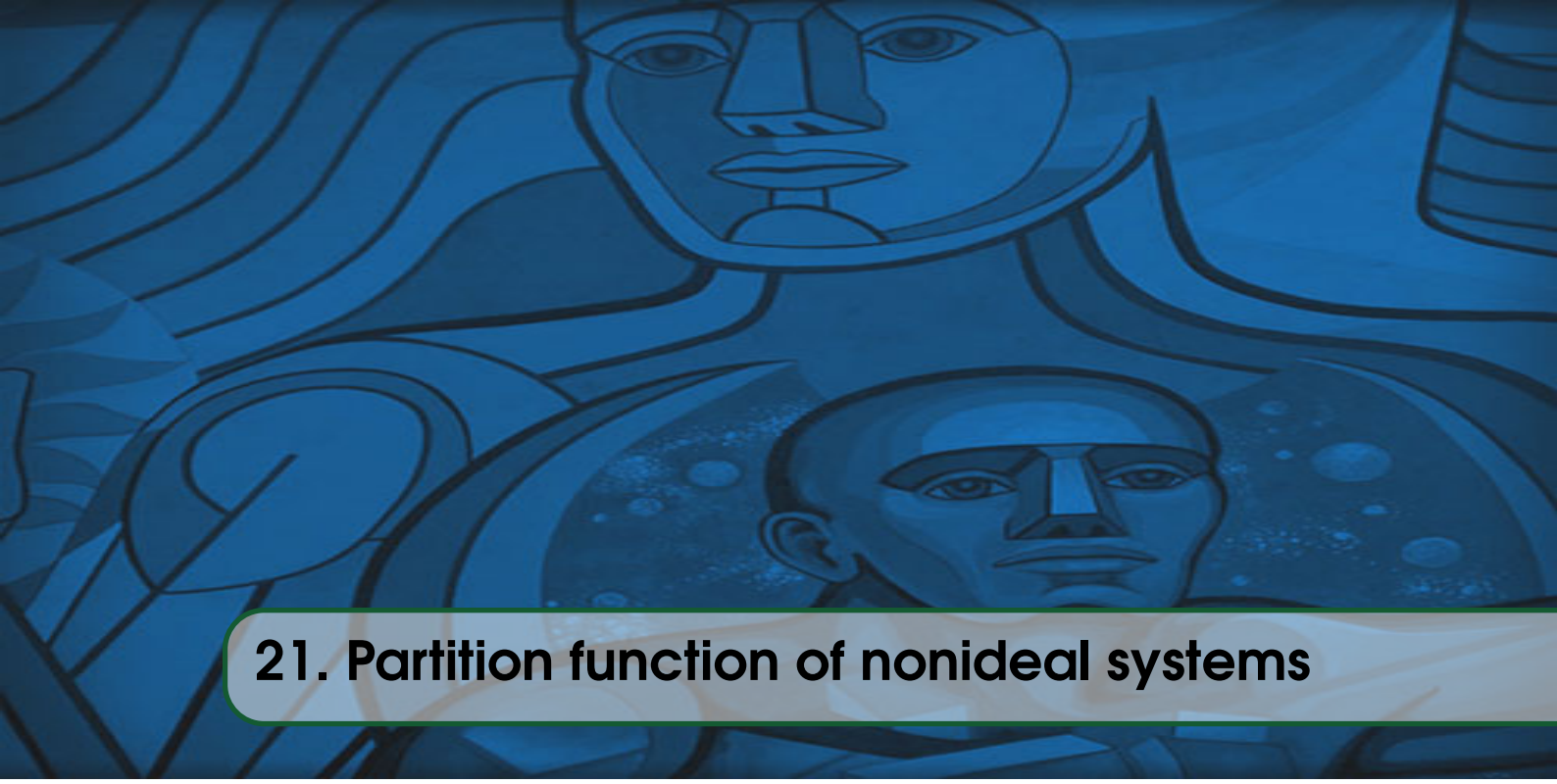
Exercise 20.1 Calculate the Laplace transform $\tilde{k}(\beta P)$ according to Eq. (20.14) for the model with the square-well square shoulder potential (it may have thermodynamic anomalies),

$$u(x) = \begin{cases} \infty, & x < a \\ \varepsilon_1, & a < x < a\lambda_1 \\ -\varepsilon_2, & a\lambda_1 < x < a\lambda_2 \\ 0, & x > a\lambda_2. \end{cases} \tag{20.27}$$

Just insert this potential into the definition of the Laplace transform and make necessary integrations. The results can be put into the expression for the chemical potential and then the equation of state can be obtained. ■

Interestingly, the primitive models explored in this chapter have been studied in context of inhomogeneous one dimensional fluids and within the grand canonical ensemble

procedure, see e.g. the works of H.T. Davis [31], P.A. Monson [33].



21. Partition function of nonideal systems

21.1 Partition function of nonideal classical system of particles

In this chapter we turn our attention to the calculation of partition function and thermodynamic properties of a classical non-ideal system of particles. The Hamiltonian of the classical one-component system we are dealing with is,

$$H(p,q) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} + U_N(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N), \tag{21.1}$$

where p_i are the Cartesian components of the momenta of particles and U_N is the potential energy of interactions,

$$U_N(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N) = \sum_{i,j=1(i<j)}^{N} u(\mathbf{q}_i, \mathbf{q}_j), \tag{21.2}$$

consisting of pair interaction potentials dependent on the coordinates of particles.

The partition function of this system within the canonical ensemble setup (N,V,T) is,

$$Q_N(V,T) = \frac{1}{N!(2\pi\hbar)^{3N}} \int ... \int \prod_{i=1}^{3N} dp_i dq_i \, e^{-\beta H(p,q)}. \tag{21.3}$$

If the system is ideal, i.e. $U_N = 0$, the partition function of a classical ideal gas is obtained straightforwardly,

$$Q_N^{(id)}(V,T) = \frac{V^N}{N!(2\pi\hbar)^{3N}} (2\pi m k_B T)^{3N/2}. \tag{21.4}$$

It can be conveniently expressed in terms of the de Broglie wave length, $\lambda = h/(2\pi m k_B T)^{1/2}$, $Q_N^{(id)}(V,T) = (V/\lambda^3)^N/N!$. Then, all thermodynamic properties follow from the Helmholtz free energy, $A = -k_B T \ln Q_N^{(id)}(V,T)$.

Our interest is, however, in the nonideal system of interacting particles in which $U_N \neq 0$. In this case exact expression for the partition function cannot be obtained and one needs to resort to approximate methods of calculation of the configurational partition function, Q_N defined as,

$$Q_N(V,T) = Q_N^{(id)}(V,T) Z_N, \tag{21.5}$$

where,

$$Z_N = \frac{1}{V^N} \int_V ... \int_V (\prod_{i=1}^{N} d^3 q_i)\ e^{-\beta U_N(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N)}. \tag{21.6}$$

Most easily, the results can be achieved, if one restricts the study to a dilute gas. To be specific, let us consider a system at rather high temperature to use Boltzmann statistics for a classical system. In addition, suppose that the average distance between particles is much larger than the radius of action of interparticle interactions, r_0,

$$< r > >> r_0. \tag{21.7}$$

The average inter-particle distance is related to the gas density, $n = N/V$,

$$\frac{4\pi}{3} < r >^3 = \frac{V}{N} = \frac{1}{n}. \tag{21.8}$$

In order to justify the application of classical statistics, we would like to provide estimates of the characteristics of real gases in Table 21.1. The triple point temperature, $T_{tr}(K)$ (it is the temperature where three phases, solid, liquid and gas coexist) is presented to give an idea how far is the room temperature from such conditions, next the de Broglie thermal wave length is calculated and the average inter-particle separation in a gas is evaluated from the experimental density.

Table 21.1: Characteristics of noble gases.

System	$T_{tr}(K)$	λ (Å)	$\lambda / < r >$
Ne	24.4	0.78	0.26
Ar	84	0.30	0.083
Kr	116.6	0.176	0.046
Xe	161.3	0.120	0.029

It can be seen that except Ne, classical treatment is entirely satisfactory and the quantum corrections are small.

If the system in question is dilute, we deal with the case,

$$n << \frac{3}{4\pi r_0^3}, \tag{21.9}$$

For illustrative purposes, recall the form of the Lennard-Jones (12-6) inter-particle interaction from chapter 6,

$$u(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right], \tag{21.10}$$

where σ and ε are the parameters of the potential that describe the diameter of particles and the strength of pair interaction. The parameter σ corresponds to $u(\sigma) = 0$ whereas the minimum of the potential is at $r_{min} = 2^{1/6}\sigma$.

The Lennard-Jones potential is well suited to describe the properties of the noble gases referred also as the inert gases. Various sets of parameters can be found in scientific literature, see the Table 21.2 below and the following reference for a more detailed description: How well does the Lennard-Jones potential represent the thermodynamic properties of noble gases?, by G. Rutkai *et al.*, [34].

Usually, the effects of the long-range tail of the potential are neglected. In other words, certain cut-off distance is assumed while developing theoretical constructions or applying computer simulation procedures. Common cut-off distance value is at 3σ or say at

Table 21.2: Parameters of the Lennard-Jones potential for noble gases.

Gas	σ (nm)	ε/k_B (K)
He	0.2628	5.465
Ne	0.2775	36.831
Ar	0.3401	116.81
Kr	0.3601	164.56
Xe	0.4055	218.18

≈ 10Å. Consequently, the condition of dilute system given by Eq. (21.9) is $n << 10^{21} cm^{-3}$. However, real gases at normal conditions are characterized by the density of the order $\approx 10^{19} cm^{-3}$ (this estimate can be easily obtained by taking tables of density of gases and their molar mass values). Thus real gases satisfy the definition of the system being dilute.

In order to get insight into physical consequences of this situation, we would like to make the following comments. The probability that two particles interact between themselves, is in fact the probability that we find them inside the sphere determined by the radius of interaction potential. If one neglects the influence of a particle on the coordinate of the other one, i.e. two particles are independent, the desired probability would become the product of finding both particles in the sphere with the volume $4\pi < r_0 >^3$ with respect to the volume each of the particles occupy in the system, i.e. w.r.t. $4\pi < r >^3$. Thus, it is proportional to $(r_0^3 / < r >^3)^2$ for pair interaction or say for possible pair collisions. Along this line of thinking we can claim that the probability that three particles interact or "collide" simultaneously is proportional to $(r_0^3 / < r >^3)^3$. However, in a dilute regime, $(r_0^3 / < r >^3)^2 >> (r_0^3 / < r >^3)^3$ evidently. Hence, while calculating the configurational partition function, we will neglect the terms describing configurations with triplets of closely situated particles and higher order terms.

Proceed to the explicit evaluation of Q_N by using Eq. (21.6). To do that, explore the integrand of Eq. (21.6), and write it as follows,

$$\exp\{-\beta U_N(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N)\} = \exp\left\{-\beta \sum_{i,j=1(i<j)}^{N} u(\mathbf{q}_i, \mathbf{q}_j)\right\} = \prod_{1 \le i < j \le N} e^{-\beta u(\mathbf{q}_i, \mathbf{q}_j)}. \tag{21.11}$$

The last equality in the r.h.s. of this equation is the product of $N(N-1)$ factors, each depends on the coordinates of two particles. The problem while dealing with the Boltzmann factors of the type $\exp\{-\beta u(r_{ij})\}$ is that it tends to unity as the potential tends to zero at large inter-particle separation, thus the integration to infinity (i.e. over the macroscopic volume is problematic). Therefore, for the sake of convenience introduce the so-called Mayer functions via the relation,

$$f_{ij} = f(\mathbf{q}_i, \mathbf{q}_j) = \exp\{-\beta[u(\mathbf{q}_i, \mathbf{q}_j)]\} - 1. \tag{21.12}$$

The Mayer functions are finite at all inter-particle distances and tend to zero at infinite separation of two particles (see Fig. 21.1).

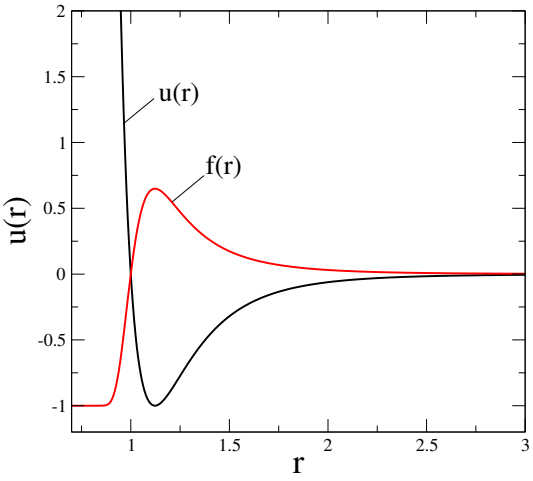


Fig. 21.1: The interaction potential and the corresponding Mayer function.

Then, the configurational partition function is,

$$Z_N = \frac{1}{V^N} \int_V ... \int_V \left(\prod_{i=1}^{N} d^3 q_i\right) \prod_{i<j}^{N} (1 + f_{ij}). \tag{21.13}$$

To be more explicit, this equation is,

$$Z_N = \frac{1}{V^N} \int_V ... \int_V d\mathbf{q}_1 d\mathbf{q}_2 ... d\mathbf{q}_N \{1 + \sum_{i<j} f_{ij} + \sum\sum f_{ij} f_{kn} + \sum\sum\sum f_{ij} f_{kn} f_{rt} + ...\}, \tag{21.14}$$

where the integrand is the series of terms each containing products of sums with increasing degree of complexity. Each of the terms in the r.h.s. can be integrated over the variables that are not involved in the indices of the Mayer functions to yield,

$$Z_N = 1 + \frac{1}{V^2} \sum_{i<j} \int_V \int_V d\mathbf{q}_i d\mathbf{q}_j f_{ij} + \frac{1}{V^4} \sum_{i<j} \sum_{k<n} \int_V \int_V \int_V \int_V d\mathbf{q}_i d\mathbf{q}_j d\mathbf{q}_k d\mathbf{q}_n f_{ij} f_{kn} + ... \tag{21.15}$$

Let us consider the most simple integral with a single Mayer function,

$$\int_V \int_V d\mathbf{q}_i d\mathbf{q}_j f_{ij} = \int_V \int_V d^3 q_i d^3 q_j f(|\mathbf{q}_i - \mathbf{q}_j|). \tag{21.16}$$

The Mayer function depends on the absolute value of the difference between coordinates only because the system is homogeneous and isotropic. Moreover, the integral does not depend on the particles' numbers because the Mayer function is the same for any pair of particles.

Let us make a coordinate transformation, $\mathbf{q}_i = \mathbf{q}' + \mathbf{q}$ and $\mathbf{q}_j = \mathbf{q}$. Jacobian of such transformation is unity. Then,

$$\int_V \int_V d^3 q_i d^3 q_j f(|\mathbf{q}_i - \mathbf{q}_j|) = V \int_V d^3 q' f(|\mathbf{q}'|). \tag{21.17}$$

We can introduce notation,

$$\beta_1 = \int_V d^3 q f(|\mathbf{q}|) = \int_V d^3 q (e^{-u(|\mathbf{q}|)/k_B T} - 1) \tag{21.18}$$

called as the first irreducible integral. Because the Mayer function is practically zero for inter-particle distances larger than the range of the potential, r_0, the integral β_1 is the

function of temperature but does not depend on the volume. Returning to Eq. (21.15), we observe that the number of such types of terms yielding β_1 is $N(N-1)/2$.

Next, consider the integral that involves a product of two Mayer functions, cf. Eq. (21.15),

$$\int\limits_V \int\limits_V \int\limits_V \int\limits_V d\mathbf{q}_i d\mathbf{q}_j d\mathbf{q}_k d\mathbf{q}_n f_{ij} f_{kn} \tag{21.19}$$

If there are no coinciding indices between two pairs, the integration will lead to the term $V^2\beta_1^2$. There exists a possibility that one of the indices in two pairs coincides, e.g. $j = k$ or $j = n$ or $i = n$ (recall that $f_{ij} = f_{ji}$). Then, the term above reduces to,

$$V \int\limits_V \int\limits_V \int\limits_V d\mathbf{q}_i d\mathbf{q}_j d\mathbf{q}_n f_{ij} f_{jn}. \tag{21.20}$$

Let apply, change of coordinates such that $\mathbf{q}_j$ is at the origin of coordinate system, e.g.

$$\mathbf{q}_i = \mathbf{q} + \mathbf{q}_j; \quad \mathbf{q}_n = \mathbf{q}^{'} + \mathbf{q}_j; \quad \mathbf{q}_i = \mathbf{q}_j. \tag{21.21}$$

The Jacobian of this linear transformation is unity. Then,

$$\begin{aligned} V \int\limits_V \int\limits_V \int\limits_V d^3q_i d^3q_j d^3q_n f_{ij} f_{jn} &= V \int\limits_V \int\limits_V \int\limits_V d^3q d^3q^{'} d^3q_j f(|\mathbf{q}|) f(|q^{'}|), \\ &= V^2 \int\limits_V \int\limits_V d^3q d^3q^{'} f(|\mathbf{q}|) f(|q^{'}|). \end{aligned} \tag{21.22}$$

Thus, we obtain $V^2\beta_1^2$ again.

Next order term in Eq. (21.9) contains the product of three Mayer functions. Then, it can happen that some integrals cannot be reduced to the product of β_1. As an example, consider the following term,

$$\int\limits_V \int\limits_V \int\limits_V d^3q_1 d^3q_2 d^3q_3 f_{12} f_{13} f_{23}. \tag{21.23}$$

This integral will be non-zero only if the following inequalities are satisfied, $|\mathbf{q}_1 - \mathbf{q}_2| < r_0$, $|\mathbf{q}_1 - \mathbf{q}_3| < r_0$ and $|\mathbf{q}_2 - \mathbf{q}_3| < r_0$, i.e. when three particles collide or in other words are situated inside the sphere of radius r_0 simultaneously. However, as we discussed above, such events in a dilute system are rare and will be neglected. So, in fact we assume that all terms in the entire series given by Eq. (21.9) can be expressed in terms of β_1, if the system in question is dilute.

The number of equal terms for each order of the expansion is calculated straightforwardly. Namely, if we need to pick up a pair from the number of pairs $N(N-1)/2)$ (for the term with the product of two Mayer functions), the number of terms is,

$$\frac{\left[\frac{N(N-1)}{2}\right]!}{2!\left[\frac{N(N-1)}{2} - 2\right]!} = \frac{1}{2!}\left[\frac{N(N-1)}{2}\right]\left[\frac{N(N-1)}{2} - 1\right]. \tag{21.24}$$

Next, we will need to pick up three pairs of the total number of pairs (for the term with the product of three Mayer functions), i.e.,

$$\frac{\left[\frac{N(N-1)}{2}\right]!}{3!\left[\frac{N(N-1)}{2} - 3\right]!} = \frac{1}{2!}\left[\frac{N(N-1)}{2}\right]\left[\frac{N(N-1)}{2} - 1\right]\left[\frac{N(N-1)}{2} - 2\right]. \tag{21.25}$$

Recall, that thermodynamic limit, $N \to \infty$, $V \to \infty$ with $n = N/V = const$ should be performed.

Then, the configuration partition functions with the assumed approximations, attains the final form,

$$Z_N = 1 + \frac{N^2}{2V}\beta_1 + \frac{1}{2}\left(\frac{N^2}{2V}\right)^2 \beta_1^2 + \frac{1}{3!}\left(\frac{N^2}{2V}\right)^3 \beta_1^3 + ... = \exp\left(\frac{N^2}{2V}\beta_1\right). \tag{21.26}$$

Inserting this expression into the partition function given by Eq. (21.5), one obtains the Helmholtz free energy of a dilute nonideal system as,

$$A = A_{id} - Nk_BT \times \frac{N}{2V}\beta_1(T). \tag{21.27}$$

This expression leads to the equation of state,

$$P = -\left(\frac{\partial A}{\partial V}\right)\Big|_{N,T} = P_{id} - \frac{N^2 k_B T}{2V^2}\beta_1, \tag{21.28}$$

with $P_{id} = Nk_BT/V$.

What kind of physics we have obtained with this result? In order to make one step further, it is necessary to choose the model of inter-particle potential and to calculate the first irreducible integral, β_1. Consider a model potential of the form,

$$u(r) = \begin{cases} \infty, & r \le d \\ -\frac{\zeta}{r^6}, & r > d \end{cases}. \tag{21.29}$$

The model includes a hard-core repulsion and an attractive term mimicking attractive tail of the Lennard-Jones interaction potential. The strength of attraction is determined by the value of the parameter ζ. Concerning the strength of attraction we assume that,

$$\frac{|u_{min}|}{k_BT} << 1. \tag{21.30}$$

The reason of doing that is to avoid the possibility of condensation expected upon cooling the system. This restriction is consistent with the assumption that the system is dilute or say theoretical result is not designed to capture transformation of the system into dense phase.

Let us evaluate β_1 now,

$$\beta_1 = 4\pi \int_0^\infty dr r^2 (e^{-u(r)/k_BT} - 1) = -\frac{4\pi}{3}d^3 + 4\pi \int_d^\infty dr r^2 (e^{-u(r)/k_BT} - 1). \tag{21.31}$$

Our assumption from Eq. (21.29) permits to make series expansion of the Boltzmann multiplier, $\exp(-u(r)/k_BT)$, and keep two terms only. Then,

$$\beta_1 = -\frac{4\pi}{3}d^3 + \frac{4\pi}{3k_BT}\frac{\zeta}{d^3}. \tag{21.32}$$

In summary, our equation of state is,

$$P = \frac{Nk_BT}{V} + \frac{2\pi}{3}\frac{N^2}{V^2}d^3 k_BT - \frac{2\pi}{3}\frac{N^2}{V^2}\frac{\zeta}{d^3}. \tag{21.33}$$

Interestingly, this equation of state has similarity with the seminal van der Waals equation of state (vdW EOS),

$$\left(P + \frac{N^2}{V^2}a\right)(V - Nb) = Nk_BT, \tag{21.34}$$

where a and b are the parameters of the equation of state. To illustrate this similarity, let us rewrite the vdW EOS as follows,

$$P = \frac{Nk_BT}{V - Nb} - \frac{N^2}{V^2}a = \frac{Nk_BT/V}{\left(1 - \frac{N}{V}b\right)} - \frac{N^2}{V^2}a. \tag{21.35}$$

Next, we would like to adjust this result to dilute fluid system, or in other words to take into account that $Nb/V << 1$. Then Eq. (21.34) takes the form,

$$P = \frac{Nk_BT}{V} + \frac{N^2}{V^2}(k_BTb - a). \tag{21.36}$$

Thus, from the comparison of Eqs. (21.35) and (21.32), we see that,

$$b = \frac{2\pi}{3}d^3, \quad a = \frac{2\pi}{3}\frac{\zeta}{d^3}. \tag{21.37}$$

Consequently, the parameter b is four times the volume of a particle. In order to establish meaning of the parameter a, let us note that,

$$2a = 4\pi \int_d^\infty drr^2 \frac{\zeta}{r^6}. \tag{21.38}$$

Therefore, a is one half of the energy of attractive interaction of the model.

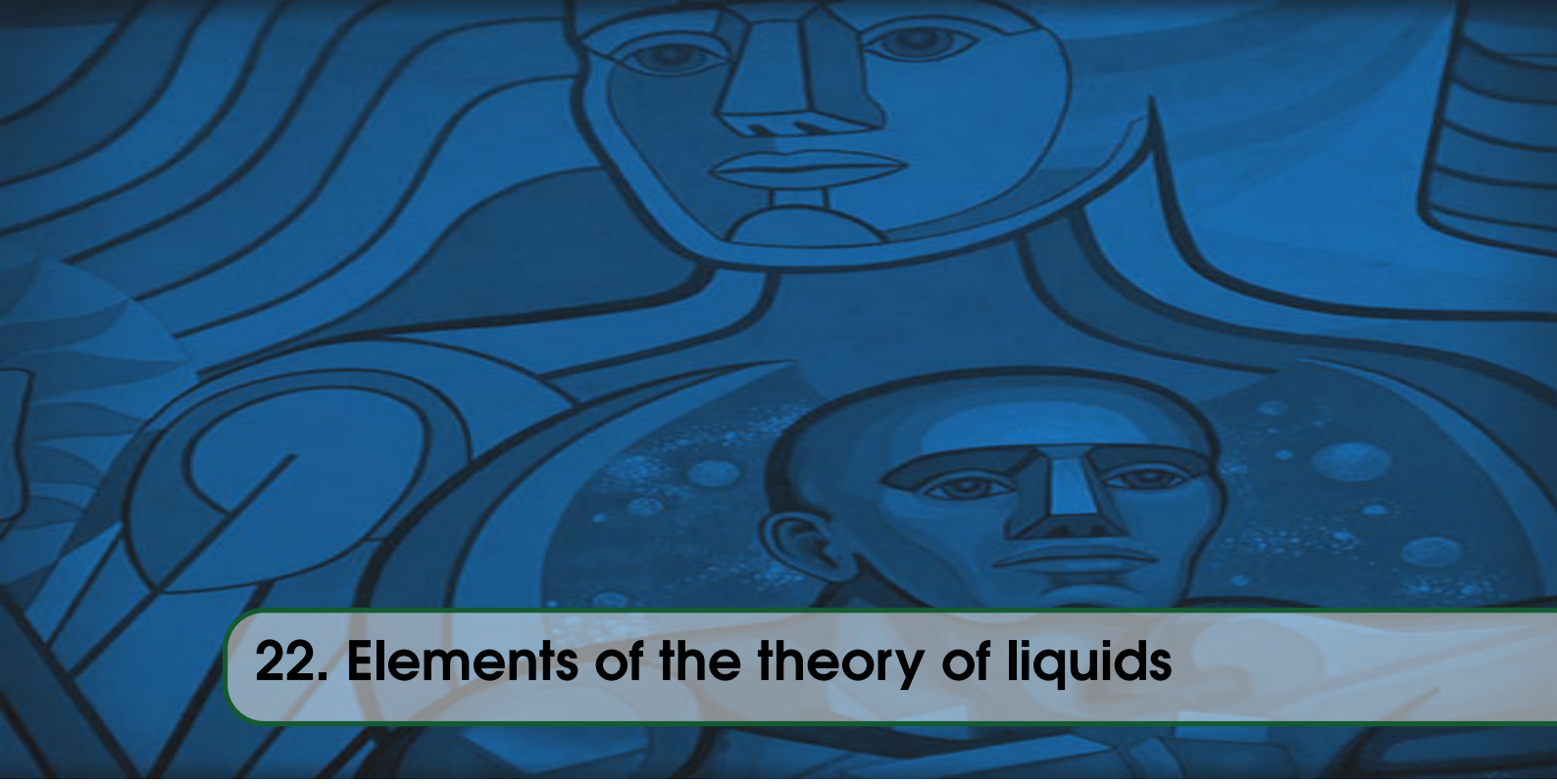
Finally, let us make summarizing comments. The starting expression for the configurational partition function given by Eq. (21.15) is exact. Next, however, for pedagogical reasons we attempted approximate calculation of Q_N by assuming that the system is dilute to obtain the free energy and arrive at the result for the equation of state given in Eq. (21.28). Here, it is important to mention that there exist more sophisticated methods to calculate the configuration partition function without resorting to the assumption of dilute system (see e.g. Terrell L. Hill, Statistical Mechanics [35]). As a result of the application of this, so-called methodology of Mayer cluster expansions, general expression for the equation of state for non-ideal system can be obtained in the form,

$$\frac{PV}{Nk_BT} = 1 + \frac{1}{V}B_2(T) + \frac{1}{V^2}B_3(T) + \dots, \tag{21.39}$$

where $B_2(T)$, $B_3(T)$, ..., are the second, the third and following virial coefficients. We can identify our result in Eq. (21.28) with the second virial coefficient $B_2(T)$. The virial coefficients of different order are reported for several substances in e.g. Refs. [7, 36]. The equation of state in the form of the virial expansion (see Eq. (21.39)) starting from Eq. (21.15) for the configuration of partition function is derived in Appendix 26.8.

However, as an attractive alternative for the calculation of the partition function, the method of the distribution functions of the theory of liquids has been developed. This issue is the subject of the following chapters.

Exercise 21.1 Using the free energy of the system, obtain the expressions for the chemical potential and entropy of the model with β_1 from Eq. (21.31). ■



22. Elements of the theory of liquids

22.1 Distribution functions in the theory of liquid state of matter

Our interest in this chapter is to introduce the concept and applications of distribution functions (sometimes they are called as the correlation functions) that describe the probability density of the coordinates of certain groups of particles with arbitrary distribution of all other particles of the system.

The probability density of the coordinates of all N particles in the system follows from the integration of the distribution function of the canonical ensemble by integration over all momenta,

$$\mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N) = \frac{1}{Z_N}\int ... \int (\prod_{i=1}^{N} d^3p_i)\exp\{-\beta E(..,p,..,q,..)\}, \tag{22.1}$$

where Z_N is the partition function of the canonical ensemble and $E(..,p,...,q,..)$ is the sum of the kinetic and potential energy of the N-particle system under study. The function $\mathcal{W}_N$ is too informative and one can search for a kind of reduced description. For this purpose, let us integrate the function above over the coordinates of all particles numbered from $s+1$ to N,

$$\mathcal{W}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = \int\limits_V ... \int\limits_V d^3q_{s+1}...d^3q_N \mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N). \tag{22.2}$$

This function provides the probability that the first particle is in the infinitesimal volume around $\mathbf{q}_1$, the second around $\mathbf{q}_2$, etc. and finally, the s-th particle at $\mathbf{q}_s$. However, the application of this kind of functions is not convenient. Namely, the probability to find one particle in the infinitesimal volume in the system with macroscopic volume V is $1/V$. Thus, the function $\mathcal{W}_s$ will be of the order $1/V^s$. Therefore, we introduce the s-order distribution function, $\mathcal{F}_s$, as follows,

$$\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = V^s\mathcal{W}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = V^s\int\limits_V ... \int\limits_V d^3q_{s+1}...d^3q_N \mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N). \tag{22.3}$$

Using this equation and the definition of $\mathcal{W}_N$ from Eq. (22.1), one obtains the normalization condition for the function of any order,

$$\frac{1}{V^s}\int\limits_V ... \int\limits_V d^3q_1...d^3q_s \mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = 1. \tag{22.4}$$

Moreover, the recursion formula between the functions of different order can be obtained. For example, consider the numbers n and s such that $n < s$. Then,

$$\frac{1}{V^n}\mathcal{F}_n(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_n) = \frac{1}{V^s}\int\limits_V ... \int\limits_V d^3q_{n+1}...d^3q_s\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s). \tag{22.5}$$

In fact, instead of these two general relations, it is sufficient to use the ones for the lowest order functions,

$$\frac{1}{V}\int\limits_V d^3q\mathcal{F}_1(\mathbf{q}) = 1, \tag{22.6}$$

and,

$$\mathcal{F}_{s-1}(\mathbf{q}_1,...,\mathbf{q}_{s-1}) = \frac{1}{V}\int\limits_V d^3q_s\mathcal{F}_s(\mathbf{q}_1,...,\mathbf{q}_s). \tag{22.7}$$

Concerning trends of behavior of the functions $F_s(\mathbf{q}_1,...,\mathbf{q}_s)$, one can deduce certain peculiarities intuitively, based on physical arguments and the properties of the probability. Specifically, let us imagine that all the particles in the system do not interact. Then, the probability to find a group of particles in certain positions would convert into the product of probabilities to find each particle in a specific position. In mathematical terms it can be written as,

$$\mathcal{W}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = \prod_{i=1}^{s}\mathcal{W}_1(\mathbf{q}_i). \tag{22.8}$$

Consequently,

$$\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = \prod_{i=1}^{s}\mathcal{F}_1(\mathbf{q}_i). \tag{22.9}$$

In reality, the particles do not interact between themselves only if they are substantially separated. Hence, the condition above in Eq. (22.9) is fulfilled if,

$$\lim_{|\mathbf{q}_i-\mathbf{q}_j|\to\infty}\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = \prod_{i=1}^{s}\mathcal{F}_1(\mathbf{q}_i), \quad i \neq j, \quad i,j = 1,2,..,s, \tag{22.10}$$

Consequently, the difference,

$$\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) - \prod_{i=1}^{s}\mathcal{F}_1(\mathbf{q}_i), \tag{22.11}$$

is the measure of correlations in the particles' coordinates due to interparticle interactions.

If the system is homogeneous, then all the coordinates of a particle are equiprobable,

$$\mathcal{W}_1(\mathbf{q}) = \frac{1}{V}, \tag{22.12}$$

and the one-particle distribution function is,

$$\mathcal{F}_1(\mathbf{q}) = 1. \tag{22.13}$$

The second-order, or the pair distribution function of the system in which the inter-particle potential depends only on the distance between particles is,

$$\mathcal{F}_2(\mathbf{q}_1,\mathbf{q}_2) = \mathcal{F}_2(|\mathbf{q}_1 - \mathbf{q}_2|). \tag{22.14}$$

This function is of primordial importance for the description of properties of homogeneous fluids.

To show that, consider the potential energy of the system,

$$U_N(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N) = \sum_{i,j=1(i<j)}^{N} u(|\mathbf{q}_i - \mathbf{q}_j|), \tag{22.15}$$

and calculate its average value,

$$\begin{aligned} < U_N > &= \frac{1}{Z_N} \int ... \int (\prod_{i=1}^{N} d^3 p_i)(\prod_{i=1}^{N} d^3 q_i) \sum_{i<j} u(q_i, q_j) e^{-\beta E(p,q)}, \\ &= \int ... \int (\prod_{i=1}^{N} d^3 q_i) \sum_{i<j} u(|\mathbf{q}_i - \mathbf{q}_j|) \mathcal{W}_N(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N), \\ &= \frac{1}{V^2} \sum_{1 \le i < j \le N} \int_V \int_V d^3 q_i d^3 q_j u(|\mathbf{q}_i - \mathbf{q}_i|) \mathcal{F}_2(|\mathbf{q}_i - \mathbf{q}_j|). \end{aligned} \tag{22.16}$$

All the integrals under the sign of sum are equal, therefore,

$$< U_N >= \frac{N(N-1)}{V^2} \int_V \int_V d^3 q_i d^3 q_j u(|\mathbf{q}_i - \mathbf{q}_i|) \mathcal{F}_2(|\mathbf{q}_i - \mathbf{q}_j|). \tag{22.17}$$

This equation permits further simplification. Namely, by transformation of coordinates (cf. previous Lecture 20) one obtains,

$$< U_N >= \frac{N(N-1)}{V} \int_V d^3 q u(|\mathbf{q}|) \mathcal{F}_2(|\mathbf{q}|). \tag{22.18}$$

Next passing to spherical coordinates and performing thermodynamic limit, the final form of the result is,

$$< U_N >= 2\pi \frac{N^2}{V} \int_0^\infty dr r^2 u(r) \mathcal{F}_2(r). \tag{22.19}$$

The left hand side is the internal energy of the system and the equation above is called as the energy equation. It expresses the internal energy of one-component fluid in terms of the interaction potential and the pair distribution function.

It seems profitable to obtain the free energy in terms of the pair distribution function rather than the internal energy. To do that, apply the so-called "charging" procedure. Let consider a system in which the potential energy U_N is substituted by the potential energy λU_N, where λ is a dimensionless parameter. Then the configurational partition function Z_N changes into $Z_N(\lambda)$ such that,

$$A(\lambda) = A_{id} - k_B T \ln Z_N(\lambda) = A_{id} - k_B T \ln \{ \frac{1}{V^N} \int_V ... \int_V (\prod_{i=1}^{N} d^3 q_i) e^{-\beta \lambda U_N} \}. \tag{22.20}$$

The derivative of the free energy by λ is,

$$\frac{\partial A(\lambda)}{\partial \lambda} = \frac{1}{V^N Z_N(\lambda)} \int_V ... \int_V (\prod_{i=1}^{N} d^3 q_i)\ U_N e^{-\beta \lambda U_N}. \tag{22.21}$$

The r.h.s. of the equation is in fact the average potential energy calculated with the distribution $\exp(-\beta \lambda U_N)$,

$$< U_N >_\lambda = \frac{1}{V^N Z_N(\lambda)} \int_V ... \int_V (\prod_{i=1}^{N} d^3 q_i)\ U_N e^{-\beta \lambda U_N}, \tag{22.22}$$

Thus, Eq. (22.21) is,

$$\frac{\partial A(\lambda)}{\partial \lambda} = < U_N >_\lambda . \tag{22.23}$$

One should take into account the boundary conditions for this differential equation,

$$A(\lambda)|_{\lambda=0} = A_{id}; \quad A(\lambda)|_{\lambda=1} = A. \tag{22.24}$$

Then, integrate the differential Eq. (22.23) to get,

$$A = A_{id} + \int_{0^+}^{1} d\lambda < U_N >_\lambda . \tag{22.25}$$

However, the internal energy of the system follows from Eq. (22.19). Therefore, the free energy can be written as,

$$A = A_{id} + 2\pi \frac{N^2}{V} \int_{0^+}^{1} d\lambda \int_0^\infty dr r^2 u(r) \mathcal{F}_2(r;\lambda), \tag{22.26}$$

where the pair distribution function $\mathcal{F}_2(r;\lambda)$ is for the system with the potential energy λU_N. In conclusion, the calculation of the free energy of a given system requires calculation of the pair distribution function for a set of systems with different degree of switching on the inter-particle interactions to perform integration over λ.

The final part of this chapter is dedicated to the calculation of the equation of state of a fluid in terms of the pair distribution function. The pressure follows from the Helmholtz free energy,

$$P = -\left(\frac{\partial A}{\partial V}\right)\Big|_{N,T} = P_{id} + k_B T \left(\frac{\partial \ln Z_N}{\partial V}\right), \tag{22.27}$$

with the configurational partition function,

$$Z_N = \frac{1}{V^N} \int_V ... \int_V (\prod_{i=1}^{N} d^3 q_i) \ \exp\{-\beta \sum_{i<j} u(|\mathbf{q}_i - \mathbf{q}_j|)\}. \tag{22.28}$$

It is worth noting that the volume change can be induced by the rescaling multiplier, i.e. instead of q let us use the coordinate $\tilde{q} = \xi q$. Then, $\tilde{V} = \xi^3 V$, $d^3\tilde{q} = \xi^3 d^3 q$. The partition function of the system with rescaled coordinates, $Z_N(\xi)$, is,

$$\begin{aligned} Z_N(\xi) &= \frac{1}{\tilde{V}^N} \int_{\tilde{V}} ... \int_{\tilde{V}} (\prod_{i=1}^{N} d^3 \tilde{q}_i) \exp\{-\beta \sum_{i<j} u(|\tilde{\mathbf{q}}_i - \tilde{\mathbf{q}}_j|)\}, \\ &= \frac{1}{V^N} \int_{\tilde{V}} ... \int_{\tilde{V}} (\prod_{i=1}^{N} d^3 q_i) \exp\{-\beta \sum_{i<j} u(\xi|\mathbf{q}_i - \mathbf{q}_j|)\}. \end{aligned} \tag{22.29}$$

The initial value of the configurational partition function would be recovered by putting $\xi = 1$ at the end of the procedure. In other words what we will do is to use the expression given by Eq. (22.29) as follows,

$$P = P_{id} + k_B T \left[\frac{1}{Z_N(\xi)} \left(\frac{\partial Z_N(\xi)}{\partial \tilde{V}}\right)\right]\Big|_{\xi=1}, \tag{22.30}$$

where

$$P = P_{id} + k_B T \frac{1}{Z_N} \left[\left(\frac{\partial Z_N(\xi)}{\partial \xi}\right)\left(\frac{d\xi}{d\tilde{V}}\right)\right]\Big|_{\xi=1}. \tag{22.31}$$

Because, $d\xi/d\tilde{V} = \xi/3\tilde{V}$, then

$$P = P_{id} + \frac{k_B T}{Z_N}\frac{1}{3V}\left(\frac{\partial Z_N(\xi)}{\partial \xi}\right)\Big|_{\xi=1}. \tag{22.32}$$

Let us calculate the auxiliary derivative,

$$\begin{aligned}\frac{\partial Z_N(\xi)}{\partial \xi} &= \frac{1}{V^N}\int_{\tilde{V}}\dots\int_{\tilde{V}}(\prod_{i=1}^{N} d^3q_i)\frac{\partial}{\partial \xi}\exp\{-\beta\sum_{i<j}u(\xi|\mathbf{q}_i-\mathbf{q}_j|)\}\\ &= \left(\frac{-\beta}{V^N}\right)\sum_{1\le i<j\le N}\int_{\tilde{V}}\dots\int_{\tilde{V}}(\prod_{i=1}^{N} d^3q_i)(|\mathbf{q}_i-\mathbf{q}_j|)\frac{u(\xi|\mathbf{q}_i-\mathbf{q}_j|)}{\partial \xi}\exp\{-\beta\sum_{i<j}u(\xi|\mathbf{q}_i-\mathbf{q}_j|)\}\\ &= \left(\frac{-\beta N(N-1)}{2V^N}\right)\int_{\tilde{V}}\dots\int_{\tilde{V}}(\prod_{i=1}^{N} d^3q_i)(|\mathbf{q}_1-\mathbf{q}_2|)\frac{\partial u(\xi|\mathbf{q}_1-\mathbf{q}_2|)}{\partial(\xi|\mathbf{q}_1-\mathbf{q}_2|)}\exp\{-\beta\sum_{i<j}u(\xi|\mathbf{q}_i-\mathbf{q}_j|)\}.\end{aligned} \tag{22.33}$$

With this result, the pressure becomes,

$$P = P_{id} - \frac{N(N-1)}{2V^N Z_N}\frac{1}{3V}\int_{V}\dots\int_{V}(\prod_{i=1}^{N} d^3q_i)(|\mathbf{q}_1-\mathbf{q}_2|)u'(\mathbf{q}_1-\mathbf{q}_2|)e^{-\beta U_N}, \tag{22.34}$$

where

$$\left(\frac{\partial u(\xi|\mathbf{q}_1-\mathbf{q}_2|)}{\partial(\xi|\mathbf{q}_1-\mathbf{q}_2|)}\right)\Big|_{\xi=1} = u'(\mathbf{q}_1-\mathbf{q}_2|). \tag{22.35}$$

However,

$$\frac{1}{V^N Z_N}\int_{V}\dots\int_{V} d^3q_3\dots d^3q_N e^{-\beta U_N} = \frac{1}{V^2}\mathcal{F}_2(\mathbf{q}_1,\mathbf{q}_2). \tag{22.36}$$

Thus,

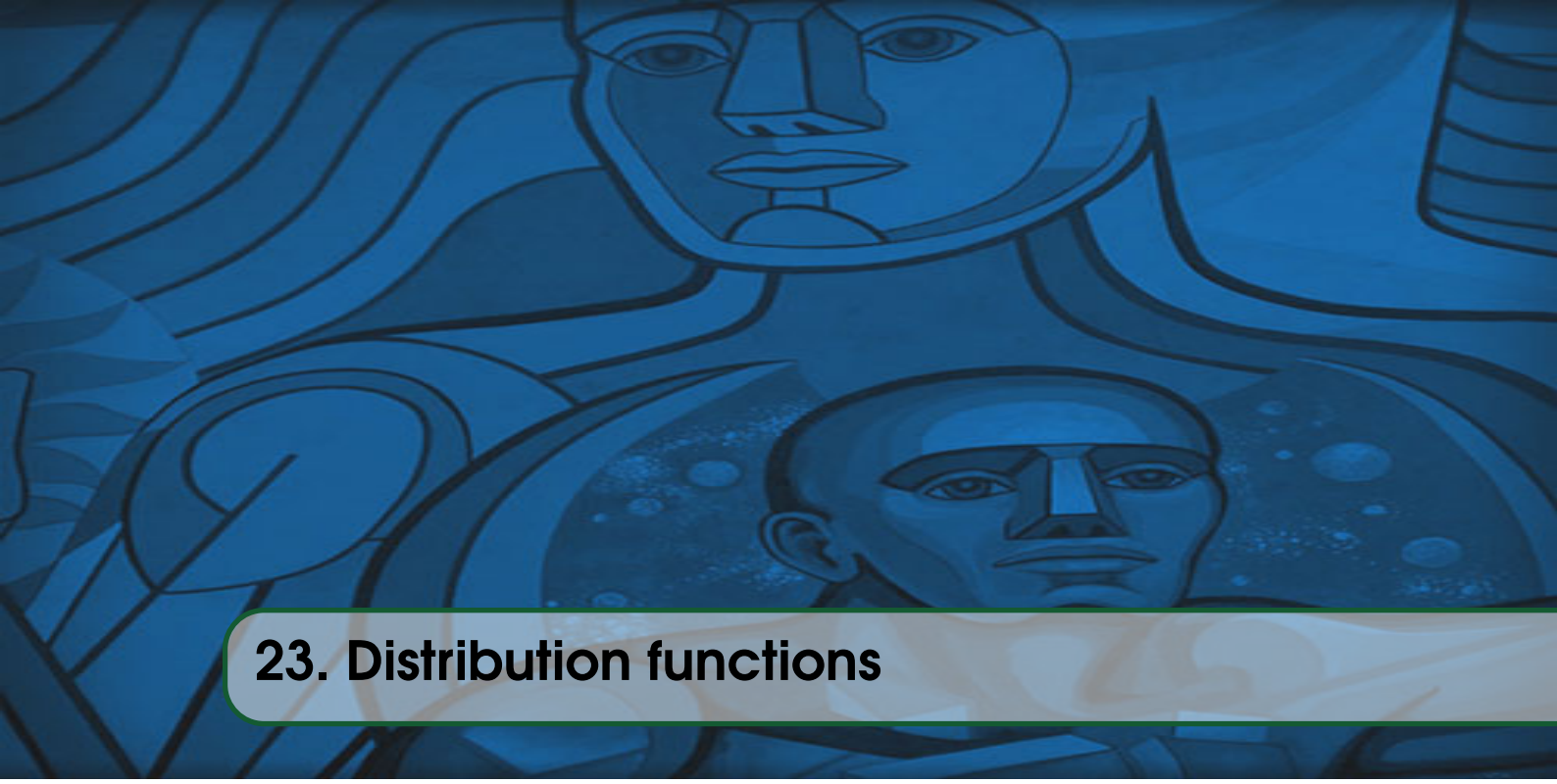
$$P = P_{id} - \frac{N^2}{2V^2}\frac{1}{3V}\int_{V}\int_{V} d^3q_1 d^3q_2(|\mathbf{q}_1-\mathbf{q}_2|)u'(\mathbf{q}_1-\mathbf{q}_2|)\mathcal{F}_2(\mathbf{q}_1,\mathbf{q}_2). \tag{22.37}$$

The system is homogeneous, consequently we introduce the notation $\mathcal{F}_2(\mathbf{q}_1,\mathbf{q}_2) = \mathcal{F}_2(|\mathbf{q}_1 - \mathbf{q}_2|) = \mathcal{F}_2(r)$, and make transformation to spherical coordinates to obtain,

$$P = P_{id} - 2\pi\frac{N^2}{3V^2}\int_0^{\infty} dr r^3 u'(r)\mathcal{F}_2(r). \tag{22.38}$$

We observe, that the pressure depends on the derivative of the interaction potential rather than on the potential itself as it is in the energy equation. The equation for pressure given above is called the virial equation of state as it involves the force expressed by the derivative of the potential multiplied by the inter-particle distance, this product is "weighted" by the pair distribution function describing average over the system's volume.

Utility of the results of this chapter are intrinsically related with the ability to calculate the pair distribution function, this issue is considered in the following chapter.



23. Distribution functions

23.1 Integro-differential equation for the distribution functions

Utility of the concept of distribution functions is in fact determined by the ability to determine them. This task is formally equivalent to the calculation of the configurational partition function but possibly may be somewhat easier.

We will start our development from the definition of the distribution function,

$$\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s) = V^s \int_V ... \int_V d^3q_{s+1}...d^3q_N \mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N). \tag{23.1}$$

Then, let us take derivative of this equation by the coordinate of any particle that is fixed on the l.h.s. for example by the component α of $\mathbf{q}_1$,

$$\frac{\partial \mathcal{F}_s}{\partial q_1^{(\alpha)}} = -\frac{V^s}{k_B T} \int_V ... \int_V d^3q_{s+1}...d^3q_N \left(\frac{\partial U_N}{\partial q_1^{(\alpha)}} \right) \mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N). \tag{23.2}$$

The potential energy of the entire system of particles,

$$U_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N) = \sum_{1\le i<j\le N} u(\mathbf{q}_i,\mathbf{q}_j), \tag{23.3}$$

can be written as the sum of three contributions to separate the groups of particles involving $\mathbf{q}_1$, namely,

$$U_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N) = \sum_{1<j\le s} u(\mathbf{q}_1,\mathbf{q}_j) + \sum_{s<j\le N} u(\mathbf{q}_1,\mathbf{q}_j) + \sum_{2\le i<j\le N} u(\mathbf{q}_i,\mathbf{q}_j). \tag{23.4}$$

Only the first two terms are involved in the differentiation in Eq. (23.2). Thus,

$$\frac{\partial U_N}{\partial q_1^{(\alpha)}} = \sum_{1<j\le s} \frac{\partial u(\mathbf{q}_1,\mathbf{q}_j)}{\partial q_1^{(\alpha)}} + \sum_{s<j\le N} \frac{\partial u(\mathbf{q}_1,\mathbf{q}_j)}{\partial q_1^{(\alpha)}}. \tag{23.5}$$

Consequently, Eq. (23.2) becomes,

$$\begin{aligned} \frac{\partial \mathcal{F}_s}{\partial q_1^{(\alpha)}} &= -\frac{V^s}{k_B T} \sum_{1<j\le s} \frac{\partial u(\mathbf{q}_1,\mathbf{q}_j)}{\partial q_1^{(\alpha)}} \int_V ... \int_V d^3q_{s+1}...d^3q_N \mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N), \\ &= -\frac{V^s}{k_B T} \sum_{s<j\le N} \int_V d^3q_j \frac{\partial u(\mathbf{q}_1,\mathbf{q}_j)}{\partial q_1^{(\alpha)}} \int_V ... \int_V \prod_{l>s,l\ne j} d^3q_l \mathcal{W}_N(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N). \end{aligned} \tag{23.6}$$

Using the definition for the distribution function, one obtains integro-differential equation,

$$\begin{aligned}&\frac{\partial}{\partial q_1^{(\alpha)}}\mathcal{F}_s(\mathbf{q}_1,...,\mathbf{q}_s)+\frac{1}{k_BT}\frac{\partial U_s}{\partial q_1^{(\alpha)}}\mathcal{F}_s(\mathbf{q}_1,...,\mathbf{q}_s)\\&+\frac{(N-s)}{Vk_BT}\int\limits_V d^3q_{s+1}\frac{\partial u(\mathbf{q}_1,\mathbf{q}_{s+1})}{\partial q_1^{(\alpha)}}\mathcal{F}_{s+1}(\mathbf{q}_1,...,\mathbf{q}_{s+1})=0,\end{aligned}\tag{23.7}$$

where,

$$U_s=\sum_{1<j\leq s}u(\mathbf{q}_1,\mathbf{q}_j).\tag{23.8}$$

If there exists a limiting value for $\mathcal{F}_s(\mathbf{q}_1,...,\mathbf{q}_s)$ for the macroscopic system with $N\to\infty$, $V\to\infty$ such that $N/V=const$, then

$$\frac{\partial}{\partial q_1^{(\alpha)}}\mathcal{F}_s+\frac{1}{k_BT}\frac{\partial U_s}{\partial q_1^{(\alpha)}}\mathcal{F}_s+\frac{N}{Vk_BT}\int\limits_V d^3q_{s+1}\frac{\partial u(\mathbf{q}_1,\mathbf{q}_{s+1})}{\partial q_1^{(\alpha)}}\mathcal{F}_{s+1}=0.\tag{23.9}$$

However, the condition of decay of correlations,

$$\lim_{|\mathbf{q}_i-\mathbf{q}_j|\to\infty}\mathcal{F}_s(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_s)-\prod_{i=1}^{s}\mathcal{F}_1(\mathbf{q}_i)=0,\quad i\neq j,\quad i,j=1,2,..,s\tag{23.10}$$

should be satisfied, as we have already mentioned in previous chapter 22. The Eq. (23.9) is known as Born-Green-Yvon (BGY) equation for the distribution functions. It has been derived in 30th of the previous century but still in use in the liquid state theory.

The principal feature of the BGY equation is that it has coupling between lower order function $\mathcal{F}_s$ with a higher order function $\mathcal{F}_{s+1}$. Thus, to solve exactly such chain of equations is impossible in close similarity to impossibility to calculate exactly the configuration partition function. Therefore, one needs to resort to certain approximations. One scheme of this sort for a "dilute" system is described below.

We assume that the system can be characterized by certain small parameter. For example, in the case of a dilute gas the small parameter can be the ratio between the radius of action of interaction between particles to the average distance between them, i.e. $r_0<<(V/N)^{1/3}$. While r_0 is the property of the model, the other distance $(V/N)^{1/3}$ characterizes thermodynamic state of the system. We assume that the small parameter is η,

$$\eta=r_0^3/(V/N).\tag{23.11}$$

Let us search the solution of the BGY equation in the form of the series,

$$\mathcal{F}_s(\mathbf{q}_1,...,\mathbf{q}_s)=\mathcal{F}_s^{(0)}(\mathbf{q}_1,...,\mathbf{q}_s)+\eta\mathcal{F}_s^{(1)}(\mathbf{q}_1,...,\mathbf{q}_s)+\eta^2\mathcal{F}_s^{(2)}(\mathbf{q}_1,...,\mathbf{q}_s)+\cdots,\tag{23.12}$$

Here, the zero-order approximation should provide the principal contribution and the following term should be the corrections of the order $\propto\eta$, $\propto\eta^2$, etc. For the sake of convenience, choose r_0 as the length unit and rewrite the BGY equation in dimensionless units $x_i^{(\alpha)}=q^{(\alpha)}/r_0$,

$$\frac{\partial}{\partial x_1^{(\alpha)}}\mathcal{F}_s+\frac{1}{k_BT}\frac{\partial U_s}{\partial x_1^{(\alpha)}}\mathcal{F}_s+\frac{1}{k_BT}\eta\int\limits_V d^3x_{s+1}\frac{\partial u(\mathbf{q}_1,\mathbf{q}_{s+1})}{\partial x_1^{(\alpha)}}\mathcal{F}_{s+1}=0.\tag{23.13}$$

Apply then series like Eq. (23.12) for the functions $\mathcal{F}_s$ and $\mathcal{F}_{s+1}$ in Eq. (23.13) and collect terms $\propto\eta^0$, $\propto\eta^1$, etc. The single Eq. (23.13) converts then into an infinite system of

equations. Let us write down the first few of them,

$$\begin{aligned}
\frac{\partial}{\partial x_1^{(\alpha)}}\mathcal{F}_s^{(0)} + \frac{1}{k_BT}\frac{\partial U_s}{\partial x_1^{(\alpha)}}\mathcal{F}_s^{(0)} &= 0,\\
\frac{\partial}{\partial x_1^{(\alpha)}}\mathcal{F}_s^{(1)} + \frac{1}{k_BT}\frac{\partial U_s}{\partial x_1^{(\alpha)}}\mathcal{F}_s^{(1)} + \frac{1}{k_BT}\int\limits_V d^3x_{s+1}\frac{\partial u(\mathbf{x}_1,\mathbf{x}_{s+1})}{\partial x_1^{(\alpha)}}\mathcal{F}_{s+1}^{(0)} &= 0,\\
\frac{\partial}{\partial x_1^{(\alpha)}}\mathcal{F}_s^{(2)} + \frac{1}{k_BT}\frac{\partial U_s}{\partial x_1^{(\alpha)}}\mathcal{F}_s^{(2)} + \frac{1}{k_BT}\int\limits_V d^3x_{s+1}\frac{\partial u(\mathbf{x}_1,\mathbf{x}_{s+1})}{\partial x_1^{(\alpha)}}\mathcal{F}_{s+1}^{(1)} &= 0,
\end{aligned} \tag{23.14}$$

The first equation is independent of the following ones. Its solution can used for $\mathcal{F}_{s+1}^{(0)}$ in the second equation to obtain $\mathcal{F}_s^{(1)}$ and so on.

Let us solve the first from the system of equations,

$$k_BT\frac{\partial}{\partial x_1^{(\alpha)}}\ln\mathcal{F}_s^{(0)} = -\frac{\partial U_s}{\partial x_1^{(\alpha)}}, \tag{23.15}$$

to get,

$$\mathcal{F}_s^{(0)} = C_0 e^{-U_s/k_BT}, \tag{23.16}$$

where C_0 is the integration constant. In order to determine C_0, recall the boundary conditions,

$$\lim_{|\mathbf{x}_i-\mathbf{x}_j|\to\infty}\mathcal{F}_s(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_s) = 1, \tag{23.17}$$

and the fact that,

$$\lim_{|\mathbf{x}_i-\mathbf{x}_j|\to\infty}U_s(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_s) = 0. \tag{23.18}$$

Then, $C_0 = 1$ and the zero-order approximation for the arbitrary distribution function is

$$\mathcal{F}_s^{(0)} = e^{-U_s/k_BT}. \tag{23.19}$$

Specifically, the pair distribution function at this level of approximation is,

$$\mathcal{F}_2^{(0)}(r) = e^{-u(r)/k_BT}. \tag{23.20}$$

This is the simplest expression that permits to get thermodynamic properties of nonideal system using the energy and virial equations from the previous chapter. However, before doing that discuss briefly the triplet distribution function, $\mathcal{F}_3^{(0)}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)$ coming from Eqs. (23.15) and (23.16). Recall that $U_3 = u(x_{12}) + u(x_{13})$ according to Eq. (23.8). Then,

$$\mathcal{F}_3^{(0)}(x_1,x_2,x_3) = e^{-(u(x_{12})+u(x_{13}))/k_BT} = \mathcal{F}_2^{(0)}(x_{12})\mathcal{F}_2^{(0)}(x_{13}). \tag{23.21}$$

However, the entire procedure of the derivation of the BGY equation and consequently the Eq. (23.15) is satisfied for,

$$\tilde{U}_s(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_s) = \sum_{1\le i<j\le s}u(\mathbf{q}_i,\mathbf{q}_j), \tag{23.22}$$

as well. The boundary condition in the form of Eq. (23.18) is more consistently satisfied for $\tilde{U}_s$ because this function contains the interaction potentials between all pairs of particles included in the group of s of them. With this procedure, the triplet function is entirely symmetric,

$$\mathcal{F}_3^{(0)}(x_1,x_2,x_3) = e^{-(u(x_{12})+u(x_{13})+u(x_{23}))/k_BT} = \mathcal{F}_2^{(0)}(x_{12})\mathcal{F}_2^{(0)}(x_{13})\mathcal{F}_2^{(0)}(x_{23}). \tag{23.23}$$

This result for $\mathcal{F}_3(x_1, x_2, x_3)$ is the so-called superposition approximation derived first by Kirkwood. The higher-order corrections to the pair distribution function and to the distribution functions describing three and more particles formally can be obtained by implementing the results from Eqs. (23.20), (23.21) or (23.23) into the system of Eqs. (23.14). At this point, we are interested in the implementation of the zero-order approximation for $\mathcal{F}_2(r)$ to describe thermodynamic properties of a nonideal system of particles.

Specifically, let us use the result from Eq. (23.20) in the virial equation of state Eq. (22.37),

$$P = P_{id} - 2\pi \frac{N^2}{3V^2} \int\limits_0^\infty dr r^3 u^{'}(r) \mathcal{F}_2(r), \tag{23.24}$$

and calculate an auxiliary integral,

$$\begin{aligned} \int\limits_0^R dr r^3 u^{'}(r) e^{-\beta u(r)} &= -\beta^{-1} \int\limits_0^R r^3 d(e^{-\beta u(r)}) = -k_B T\, r^3 e^{-\beta u(r)} \Big|_0^R + 3k_B T \int\limits_0^R dr r^2 e^{-\beta u(r)}, \\ &= -k_B T r^3 \left(e^{-\beta u(r)} - 1 \right) \Big|_0^R + 3k_B T \int\limits_0^R dr r^2 (e^{-\beta u(r)} - 1). \end{aligned} \tag{23.25}$$

The first term tends to zero at $R \to \infty$, thus only the second term contributes to the integral. So, our result in the limit $R \to \infty$ is,

$$\int\limits_0^\infty dr r^3 u^{'}(r) e^{-\beta u(r)} = 3k_B T \int\limits_0^\infty dr r^2 (e^{-\beta u(r)} - 1) = \frac{3}{4\pi} k_B T \beta_1, \tag{23.26}$$

where β_1 is the first irreducible integral defined in Eq. (21.18) or Eq. (21.31). With this approximation for the pair distribution function, the equation of state then is,

$$P = P_{id} - 2\pi \frac{N^2}{3V^2} \frac{3}{4\pi} k_B T \beta_1 = P_{id} - \frac{N^2}{2V^2} k_B T \beta_1, \tag{23.27}$$

and coincides with Eq. (21.28) that has been obtained from the direct calculation of the configurational partition function.

23.2 Van der Waals equation of state

The model equation of state formulated by van der Waals provides the simplest picture of the relation between the intermolecular interaction and the qualitative aspects of thermodynamic behavior of gases and liquids. This equation has been formulated much earlier than the development of the theory of ensembles. On the other hand, it appeared to be very successful and faithfully serves for pedagogic purposes. From the pragmatic view of chemical engineering this equation served as basis for the construction of a more sophisticated semiempirical and empirical equations of state used in practice nowadays.

The vdW EOS reads,

$$\left(P + \frac{N^2}{V^2} a \right) (V - Nb) = N k_B T. \tag{23.28}$$

The a and b are the parameters of this equation. For the sake of convenience let us use molar volume, $\tilde{V} = V/(N/N_{av})$, and molar parameters, $\tilde{a} = aN_{av}^2$, $\tilde{b} = bN_{av}$,

$$P = \frac{RT}{(\tilde{V} - \tilde{b})} - \frac{\tilde{a}}{\tilde{V}^2}, \tag{23.29}$$

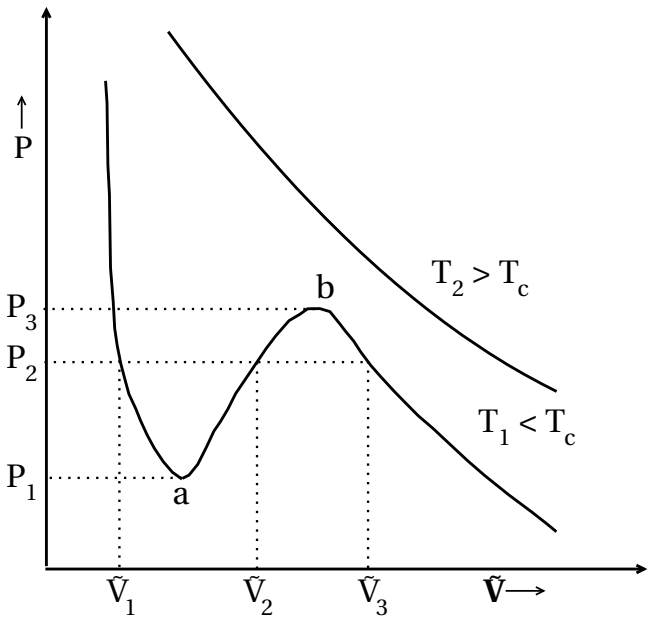


Fig. 23.1: Qualitative pressure - molar volume isotherms for van der Waals fluid. The P - $\hat{V}$ states between a and b on the T isotherm are thermodynamically unstable.

where $R = k_B N_{av}$ is the universal gas constant. We observe that the vdW EOS is the cubic equation in molar volume. In qualitative terms, the $P - \tilde{V}$ isotherms are shown in the Fig. 23.1.

At every quite arbitrary temperature above certain characteristic temperature, call it T_c, the pressure is a monotonously decreasing function of the molar volume, see e.g. the isotherm for $T = T_2$ in the figure. In other words, in this temperature range there is a one to one correspondence between pressure and density (or molar volume), only one homogeneous fluid phase exists at a given value of pressure.

In the case of a temperature below T_c, the shape of the isotherm is different. Namely, for example at $T_1 < T_c$, there exists an interval of pressures ($P_1 < P < P_3$) such that three values of molar volume correspond to every value of pressure. Thus, the EOS predicts existence of three homogeneous fluid phases. Outside of this interval of pressures, the one to one correspondence between the pressure and molar volume values preserves.

So, at T_1 and P_2, three phases with molar volumes $\tilde{V}1, \tilde{V}_2, \tilde{V}_3$ are predicted by the equation of state. However, the phase characterized by $\tilde{V}_2$ is not thermodynamically stable, because the stability requirement,

$$\left(\frac{\partial P}{\partial \tilde{V}}\right)\Bigg|_T < 0, \tag{23.30}$$

is violated for the phase and thermodynamic state with $\tilde{V}_2$. More generally, all the states between the points a and b along this isotherm are not stable from thermodynamic point of view. Evidently, the instability region shrinks and eventually disappears if temperature increases from T_1 to T_c. On the other hand, if temperature decreases starting from T_1, then the instability region becomes wider. Actually, the lower limit is at the triple point temperature, T_{tr}, mentioned in chapter 21. This issue is not discussed for the moment, however.

It is important that for a given temperature T at $T < T_c$, there exists only one value of pressure, $P(T)$, for which two stable phase with different density, or different molar volumes, coexist and are in thermodynamic equilibrium with each other. The reason is that two phases at the same pressure or say in mechanical equilibrium are not in thermodynamic equilibrium unless the chemical potential of two phases is the same, i.e. the mass balance or chemical equilibrium condition should be satisfied.

It is known from thermodynamics, that the chemical potential along an isotherm changes according to the equation (this issue will be discussed in the following chapter),

$$d\mu = \tilde{V} dP. \tag{23.31}$$

If we integrate this relation along the isotherm T_1 between two states with two respective molar volumes,$\tilde{V}1$ and $\tilde{V}_2$, then,

$$\mu_1 - \mu_2 = \int_{\tilde{V}_1}^{\tilde{V}_2} \tilde{V} dP. \tag{23.32}$$

These two phases should satisfy conditions of mechanical and chemical equilibrium, if $P_1(T_1) = P_2(T_1)$ and $\mu_1 = \mu_2$, i.e.,

$$\int_{\tilde{V}_1}^{\tilde{V}_2} \tilde{V} dP = 0. \tag{23.33}$$

This equation can be interpreted in terms of geometric construction to find two-phase equilibrium states. To do that one needs to plot a line drawn parallel to $\tilde{V}$ axis, cutting the isotherm in such a way that the area below the line till isotherm and the area above the line till isotherm are equal, see the figure. The line is called as the vapor-liquid tie line at a given temperature. By such geometric construction called the Maxwell equal area tie-line construction one can determine the entire vapor-liquid coexistence envelope or the dome shown in Fig. 23.2. It is evident from the Maxwell construction that there is only one coexistence pressure for each isotherm.

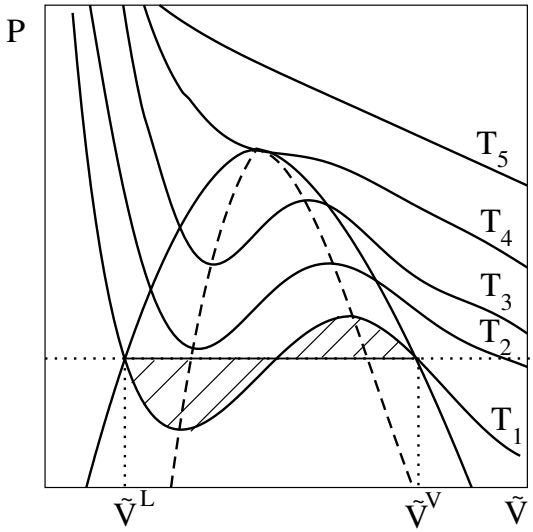


Fig. 23.2: Schematic phase diagram of a vdW fluid illustrating the vapor-liquid coexistence envelope. Maxwell equal area construction for a single isotherm is shown as well.

At the top of the dome, one can locate the critical point. It is the point with P_c, and $\tilde{V}_c$ at which the isotherm (at T_c) is tangent to the top of the two-phase coexistence. The point of inflection satisfies the conditions,

$$\left(\frac{\partial P}{\partial \tilde{V}}\right)\Big|_T = \left(\frac{\partial^2 P}{\partial \tilde{V}^2}\right)\Big|_T = 0, \quad T = T_c. \tag{23.34}$$

These conditions determine the critical pressure and critical molar volume (or equivalently the critical density). The critical parameters P_c, T_c and $\tilde{V}_c$ can be found by differentiating

the vdW EOS and using the conditions of criticality from Eq. (23.34). Namely,

$$\tilde{V}_c = 3\tilde{b}; \quad T_c = \frac{8\tilde{a}}{27R\tilde{b}}; \quad P_c = \frac{\tilde{a}}{27\tilde{b}^2}. \tag{23.35}$$

Then, one can express the parameters of the vdW EOS in terms of combinations of the critical values,

$$\tilde{a} = \frac{27R^2T_c^2}{64P_c}; \quad \tilde{b} = \frac{RT_c}{8P_c}. \tag{23.36}$$

Exercise 23.1 Derive the relations presented in Eq. (23.35) by using the vdW EOS from Eq. (23.29) together with the critical point conditions given by Eq. (23.34). Furthermore, show that,

$$\frac{P_c\tilde{V}_c}{T_c} = \frac{3R}{8}. \tag{23.37}$$

Also, show that the vdW EOS can be written in a reduced form,

$$P_r = \frac{8T_r}{3\tilde{V}_r - 1} - \frac{3}{\tilde{V}_r^2}, \tag{23.38}$$

where $P_r = P/P_c$, $T_r = T/T_c$ and $\tilde{V}_r = \tilde{V}/\tilde{V}_c$. ■

Final comments concerning the Fig. 23.2 are the following. The region labeled as metastable in this figure is located between the coexistence envelope and the unstable region. It can be considered as an "extension" of the stable parts outside of the coexistence dome. On the vapor side, the metastable region describes one-phase supercooled vapor states. Starting from the saturated vapor phase (keeping pressure constant) and moving to the left in the figure one enters metastable region of supercooled vapor, any small perturbation may break it into a two-phase, more stable, state of vapor containing liquid drops. On the opposite side starting from liquid at saturation keeping the pressure constant, one enters the metastable region by heating, any kind of perturbation may lead to a more stable two-phase state of vapor droplets dispersed in the liquid body.

The behavior of a fluid near the critical point is a wide scientific area of research. In the spirit of our discussion above, we would like to note only one important issue. The isothermal compressibility,

$$\kappa_T = -\frac{1}{\tilde{V}}\left(\frac{\partial\tilde{V}}{\partial P}\right)\Big|_T, \tag{23.39}$$

is infinite at a critical point and is very large near the critical point. The singular behavior of κ_T as well as of C_V and C_P, around the critical point are related to the large fluctuations and the developing long-range intermolecular correlations. When approaching the critical point along the critical isotherm the divergence of κ_T is of the form,

$$\kappa_T \propto (P - P_c)^{-\epsilon}, \tag{23.40}$$

where ϵ is 2/3 for the vdW EOS, and is usually from 0.75 to 0.8. On the other hand, when approaching the critical point along the critical isochore ($T > T_c$) the divergence is of the form,

$$\kappa_T \propto (T - T_c)^{-\gamma}) \propto (P - P_c)^{-\gamma}, \tag{23.41}$$

where γ is 1.0 for the vdW EOS, and is usually from 1.2 to 1.3. The isochoric heat capacity, C_V, behaves as,

$$C_V = C_0\left(\left|1 - \frac{T}{Tc}\right|\right)^{-\alpha}, \tag{23.42}$$

where $\alpha = 0.1096$, if the three-dimensional Ising model is applied.

In contrast to vapor-liquid coexistence, it is commonly believed that the transition line separating a liquid and a solid cannot be interrupted by a critical point. This opinion is based on the traditional symmetry argument that an isotropic liquid cannot be continuously transformed into a crystal with a discrete rotational and translational symmetry as documented by Landau and Lifshitz [10]. Schematic phase diagram of a simple fluid mimicking argon is shown in Fig. 23.3. In contrast to Fig. 23.2 above, here the solid phase is explicitly included.

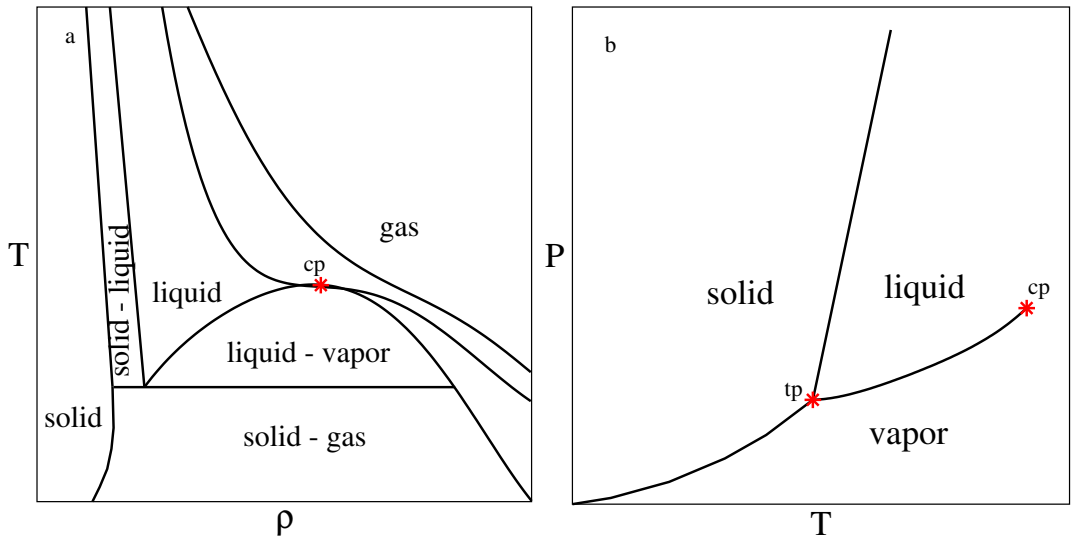


Fig. 23.3: Panel a: Schematic phase diagram of a vdW fluid illustrating the vapor-liquid coexistence envelope. The liquid-solid equilibrium is included as well. Panel b: Schematic phase diagram in $P - T$ plane.

23.3 The law of corresponding states

The phase behavior of several fluids can be reasonably well described by the vdW equation of state. Evidently, that different fluids are characterized by the different values of critical temperature, pressure and density. One may wish to perform comparison of thermodynamic properties of this kind of fluids. To do that it is profitable to introduce reduced variables,

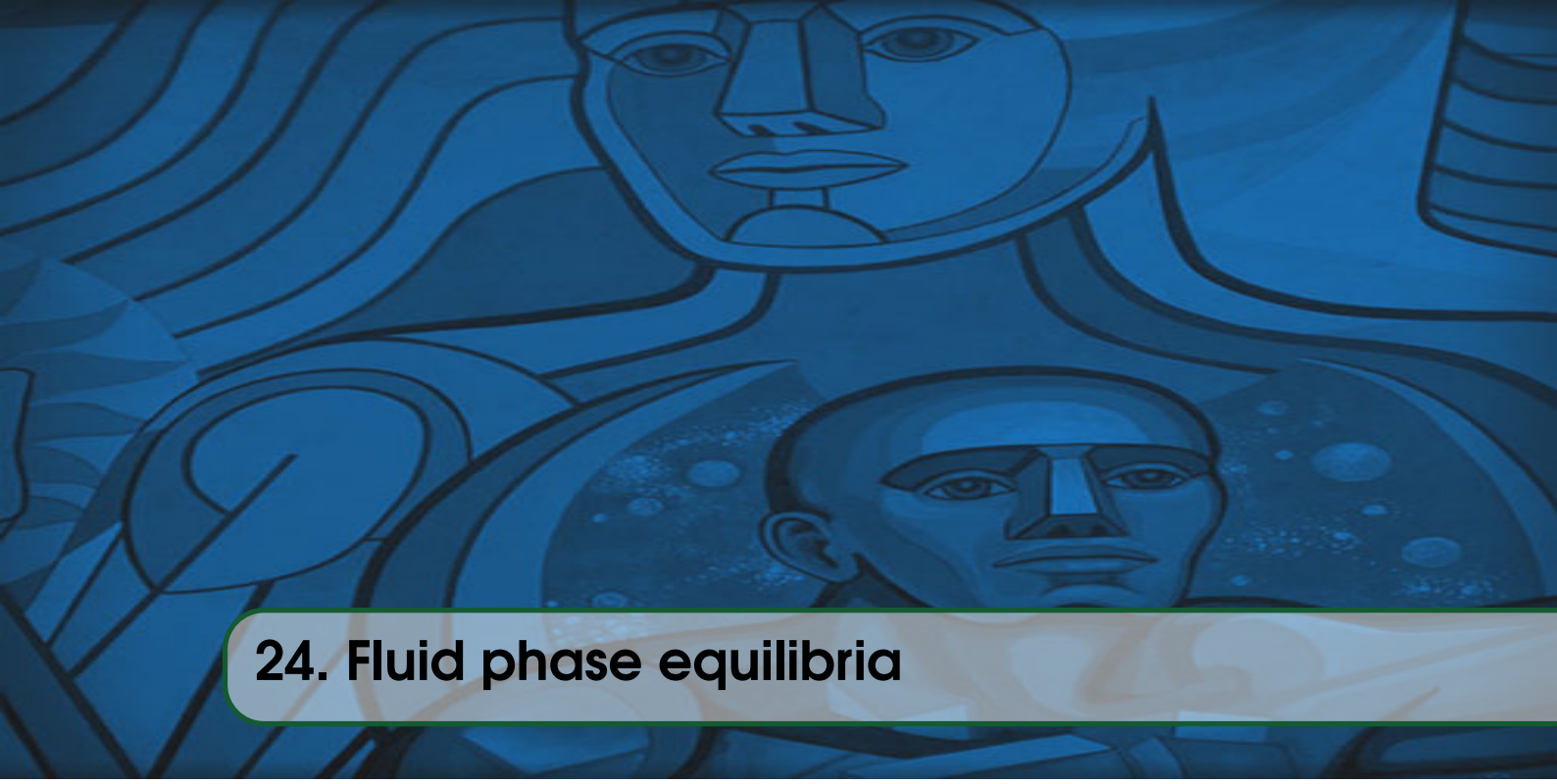
$$V_r = \frac{\tilde{V}}{\tilde{V}_c}; \quad T_r = \frac{\tilde{T}}{\tilde{T}_c}; \quad P_r = \frac{\tilde{P}}{\tilde{P}_c}, \tag{23.43}$$

If the vdW EOS is expressed in terms of these variables, one finds the "universal" equation of state,

$$P_r = \frac{8T_r}{3V_r - 1} - \frac{3}{V_r^2}. \tag{23.44}$$

This equation is the essence of the law of the corresponding states. All fluids obeying vdW equation of state will have the same reduced pressure if they are compared at the same reduced molar volume V_r and at the same reduced temperature, T_r. The law of corresponding states is obeyed quite well for a variety of fluids.

Exercise 23.2 Make a short text commenting the phase behavior shown in Fig. 23.3. Add a few comments concerning the phase diagram of water. ■



24. Fluid phase equilibria

24.1 Elements of the theory of phase equilibria

The conditions of phase equilibria follow from the second law of thermodynamics. As in the development of the theory of statistical ensembles we begin with isolated system. The second law of thermodynamics states that the entropy S of an isolated system in equilibrium is at maximum. The entropy in such setup is the function of internal energy, volume of the system and the number of particles, U, V and N. For the sake of generality, we write the last variable as N_α for each chemical component α. Assume that the system is composed of c components.

The principle of maximum entropy leads to other formulations of the second law of thermodynamics. Namely, for a closed isothermal system with a fixed volume, the Helmholtz free energy, A, $A = U - TS$ is a minimum at equilibrium. This setup corresponds to the canonical ensemble. On the other hand, if one considers a setup of the isobaric-isothermal ensemble, the second law states that the Gibbs free energy, G, $G = A + PV = H - TS$ ($H = U + PV$ is the enthalpy) attains minimum at equilibrium. Similar considerations refer to the grand canonical ensemble. In this case, the second law refers to the grand thermodynamic potential,

$$\Omega = A - \sum_{\alpha=1}^{c} \mu_\alpha N_\alpha, \tag{24.1}$$

that is at minimum for the system at equilibrium. If the first law, $dU = dQ + dW$, and the second law of thermodynamics are considered together, the fundamental equation for a homogeneous system follows,

$$dU = TdS - PdV + \sum_{\alpha=1}^{c} \mu_\alpha dN_\alpha. \tag{24.2}$$

Then,

$$\begin{aligned} dH &= TdS + VdP + \sum_{\alpha=1}^{c} \mu_\alpha dN_\alpha, \\ dG &= -SdT + VdP + \sum_{\alpha=1}^{c} \mu_\alpha dN_\alpha, \\ dA &= -SdT - PdV + \sum_{\alpha=1}^{c} \mu_\alpha dN_\alpha, \\ d\Omega &= -SdT - PdV - \sum_{\alpha=1}^{c} N_\alpha d\mu_\alpha. \end{aligned} \tag{24.3}$$

The principal issue at this point is the distinction between extensive and intensive thermodynamic variables. All of the quantities, S, U, V, N_α, A, H, G and Ω are extensive thermodynamic variables. In other words, if we consider a homogeneous system at fixed intensive variables, T, P and μ_α, and the system is increased k times in amount of each component, then the extensive variables increase in magnitude k times as well, and become: kS, kU, kV, kA, kH, kG and $k\Omega$. This particular property permits to integrate Eqs. (24.2) and (24.3). For example, perform such integration (say from x to kx) for the Gibbs free energy,

$$(k-1)G = (k-1)\sum_{\alpha=1}^{c} \mu_\alpha N_\alpha. \tag{24.4}$$

Thus, the result for a homogeneous macroscopic phase is,

$$G = \sum_{\alpha=1}^{c} \mu_\alpha N_\alpha. \tag{24.5}$$

Integration of other equations leads to the following results,

$$\begin{aligned} U &= TS - PV + \sum_{\alpha=1}^{c} \mu_\alpha N_\alpha, \\ H &= TS + \sum_{\alpha=1}^{c} \mu_\alpha N_\alpha, \\ A &= -PV + \sum_{\alpha=1}^{c} \mu_\alpha N_\alpha, \\ \Omega &= -PV. \end{aligned} \tag{24.6}$$

Intensive thermodynamic variables are constants throughout a homogeneous phase by definition and are independent of the amount of the phase.

Consider now arbitrary infinitesimal variation of the thermodynamic state of the system. Upon this change, the differential of the internal energy according to Eq. (24.6) is,

$$dU = TdS + SdT - PdV - VdP + \sum_{\alpha=1}^{c} \mu_\alpha dN_\alpha + \sum_{\alpha=1}^{c} N_\alpha d\mu_\alpha. \tag{24.7}$$

However, the fundamental Eq. (24.2) holds. Therefore, subtract Eq. (24.7) from Eq. (24.2) to obtain the condition,

$$SdT - VdP + \sum_{\alpha=1}^{c} N_\alpha d\mu_\alpha = 0. \tag{24.8}$$

This important constraint is known as the Gibbs-Duhem equation. The following conclusion then is. If we have $c+2$ intensive variables (T, P and c values for μ_α) characterizing

the homogeneous phase, only $c+1$ variables are independent. For illustration, we can say that if, for example, c values for μ_α and T are specified, then the pressure is not arbitrary, it is fixed by $P = P(T, \mu_1, ..., \mu_c)$. On the other hand, if T, P and $c-1$ chemical potentials are specified, then $\mu_\alpha = \mu_\alpha(T, P, ...\mu_{\beta\neq\alpha})$. The Gibbs-Duhem equation can be obtained by evaluation the differential for either H or A or G or Ω from Eq. (24.6) and then using fundamental Eq. (24.3).

Now turn our attention to the problem of equilibrium between phases. Evidently, thermodynamic laws should provide the solution of the problem. We will show now, how it follows from the second law of thermodynamics. Consider an isolated combined system composed of two parts, according to the setup in Fig. (24.1) below.

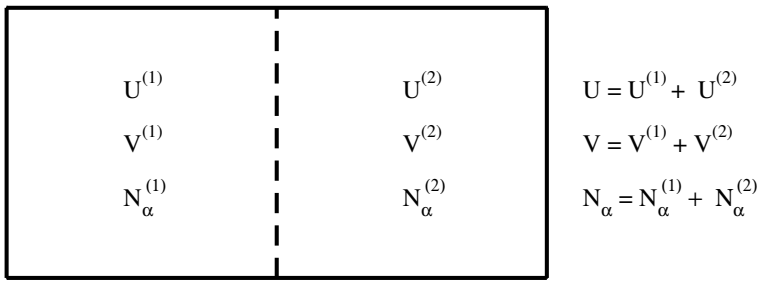


Fig. 24.1: Setup of the system to study phase equilibrium.

Suppose that the bulk homogeneous phase 1 is in equilibrium with the bulk homogeneous phase 2. The word bulk is used to explain that the phase is large enough or that it is homogeneous everywhere except at the physical interface between phases. Moreover, the entropy, energy, volume and the number of molecules within the interfacial region (evidently it exists in reality but is not considered as a separate phase) is much smaller and even negligible comparing to the values of these characteristics for each bulk phase. Consider a small change of values of the variables determining the state of phases from equilibrium, $\delta U^{(i)}, \delta V^{(i)}, \delta N_\alpha^{(i)}$, $i = 1,2$. Isolation of the entire system leads to the equations,

$$\delta U^{(1)} = -\delta U^{(2)}; \quad \delta V^{(1)} = -\delta V^{(2)}; \quad \delta N_\alpha^{(1)} = -\delta N_\alpha^{(2)}; \quad \alpha = 1,2,...,c \tag{24.9}$$

Now evaluate the change of entropy upon changing of the variables, $\delta U^{(i)}, \delta V^{(i)}, \delta N_\alpha^{(i)}$, it must obey condition,

$$\Delta S = \sum_{i=1,2} [S^{(i)}(U^{(i)} + \delta U^{(i)}, V^{(i)} + \delta V^{(i)}, \mathbf{N}^{(i)} + \delta \mathbf{N}^{(i)}) - S^{(i)}(U^{(i)}, V^{(i)}, \mathbf{N}^{(i)}] < 0, \tag{24.10}$$

and obviously should vanish at equilibrium. We would like to apply the Taylor expansion to deal with small fluctuations, thus let rewrite the inequality above as,

$$\Delta S = \sum_{k=1}^{\infty} \delta^k S < 0. \tag{24.11}$$

Here,

$$\delta S = \sum_{i=1,2} (-1)^i [(\frac{\partial S^{(i)}}{\partial U^{(i)}})\delta U^{(1)} + (\frac{\partial S^{(i)}}{\partial V^{(i)}})\delta V^{(1)} + \sum_{\alpha=1}^{c} (\frac{\partial S^{(i)}}{\partial N^{(i)_\alpha}})\delta N_\alpha^{(1)}],$$

$$\delta^2 S = \sum_{i=1,2} \frac{1}{2!} [(\frac{\partial^2 S^{(i)}}{\partial (U^{(i)})^2})(\delta U^{(1)})^2 + (\frac{\partial^2 S^{(i)}}{\partial (V^{(i)})^2})(\delta V^{(1)})^2 + \sum_{\alpha=1}^{c}\sum_{\beta=1}^{c} (\frac{\partial^2 S^{(i)}}{\partial N_\alpha^{(i)} \partial N_\beta^{(i)}})\delta N_\alpha^{(1)} \delta N_\beta^{(1)}$$

$$+2(\frac{\partial^2 S^{(i)}}{\partial U^{(i)} \partial V^{(i)}})\delta U^{(1)} \delta V^{(1)} + 2\sum_{\alpha=1}^{c} (\frac{\partial^2 S^{(i)}}{\partial U^{(i)} \partial N_\alpha^{(i)}})\delta U^{(1)} \delta N_\alpha^{(1)} + 2\sum_{\alpha=1}^{c} (\frac{\partial^2 S^{(i)}}{\partial V^{(i)} \partial N_\alpha^{(i)}})\delta V^{(1)} \delta N_\alpha^{(1)}],$$

(24.12)

and so on. From the fundamental equation for a homogeneous phase, Eq. (24.2) one has,

$$\left(\frac{\partial S^{(i)}}{\partial U^{(i)}}\right) = T^{(i)}; \quad -\left(\frac{\partial S^{(i)}}{\partial V^{(i)}}\right) = \frac{P^{(i)}}{T^{(i)}}; \quad \left(\frac{\partial S^{(i)}}{\partial N_\alpha^{(i)}}\right) = \frac{\mu_\alpha^{(i)}}{T^{(i)}}. \tag{24.13}$$

Next, choose $\delta V^{(i)} = 0$, $\delta N_\alpha^{(i)} = 0$ and $\delta U^{(i)}$ as infinitesimal fluctuation of arbitrary sign. Then,

$$\Delta S = \sum_{i=1,2} \frac{(-1)^i}{T^{(i)}} \delta U^{(1)} + O(\delta U^{(1)})^2 \tag{24.14}$$

The only way, ΔS can be less than zero for arbitrarily small deviation $\delta U^{(1)}$ is that in equilibrium state,

$$T^{(1)} = T^{(2)}. \tag{24.15}$$

Otherwise, out of equilibrium, the magnitude of $\delta U^{(1)}$ can be chosen sufficiently small to make the contribution of higher order terms negligible. The sign of $\delta U^{(1)}$ can be chosen dependent of the sign of $\sum_{i=1,2}(-1)^i/T^{(i)}$. Thus, one condition of thermodynamic equilibrium between two coexisting phases is the equality of their temperature.

Similarly, by choosing $\delta U^{(1)} = 0$, $\delta N_\alpha^{(1)} = 0$ and $\delta V^{(1)}$ infinitesimally small, one obtains,

$$\frac{P^{(1)}}{T^{(1)}} = \frac{P^{(2)}}{T^{(2)}}, \tag{24.16}$$

and finally, if $\delta U^{(1)} = 0$, $\delta V(1) = 0$, and $\delta N_\alpha^{(i)} \neq 0$ with $\delta N_{\beta\neq\alpha}^{(1)} = 0$, then,

$$\frac{\mu_\alpha^{(1)}}{T^{(1)}} = \frac{\mu_\alpha^{(2)}}{T^{(2)}}. \tag{24.17}$$

Eqs. (24.15) - (24.17) represent conditions of thermal, mechanical and chemical equilibrium of phases. Usually, the intensive thermodynamic quantities that are the same in all coexisting phases are called as field variables.

Gibbs phase rule incorporates the conditions of thermodynamic equilibrium and prescribes the number of intensive variables that can be independently set in a multiphase system at equilibrium. In order to derive the Gibbs rule suppose that there are p coexisting phases and c chemical components (a quite complex physical system by the way). The intensive variable to be specified are T, P and mole fractions $x_\alpha^{(i)}$ of each component α in each phase i. So, we have $cp + 2$ of such intensive variables. However, the chemical equilibrium requires,

$$\mu_\alpha^{(1)} = \mu_\alpha^{(2)} = ... = \mu_\alpha^{p)}, \quad \alpha = 1,2,...,c \tag{24.18}$$

From the definition of the mole fraction,

$$x_\alpha = N_\alpha / \sum_{\beta=1}^{c} N_\beta, \tag{24.19}$$

it follows that,

$$\sum_{\alpha=1}^{c} x_\alpha^{(i)} = 1, \quad i = 1,2,...,p. \tag{24.20}$$

The condition given by Eq. (24.18) yields $c(p-1)$ constraints, whereas Eq. (24.20) makes p constraints. Thus, the number f of the degrees of freedom (or intensive variables that can be set independently) is,

$$f = cp + 2 - c(p-1) - p; \quad f = 2 + c - p. \tag{24.21}$$

This equation is the prominent Gibbs rule.

For example, for a liquid-vapor equilibrium in a one-component system, $c = 1$, $p = 2$, and $f = 1$. If one intensive variable, T, is fixed then a unique liquid-vapor pair results (at a unique vapor pressure, because $P = P(T)$ along binodal). In the case of a vapor-liquid-solid equilibria $p = 3$ and then $f = 0$, thus there is a unique triple point (cf. right panel of Fig. 23.3 in chapter 23).

The equilibrium conditions we have discussed so far result from the requirement that the first variation, δS, cf. Eq. (24.12) vanishes at equilibrium. This is, however, only the necessary condition for S to be at maximum. If $\delta^2 S$ is positive for arbitrary fluctuations, $\delta U^{(i)}$, $\delta V^{(i)}$, and $\delta N_\alpha^{(i)}$, then the system is not at equilibrium and the state under consideration is not stable with respect to small fluctuations. A multiphase system will be unstable if any one of the phases is unstable to small fluctuations.

In order to get additional insights into stability conditions, let us consider a simpler case in comparison to the setup given in Fig. 24.1. Namely, assume that bulk phases 1 and 2 are identical. In this case, the system can be regarded as a single phase divided into two equal parts, labeled 1 and 2. Then, $S^{(1)} = S^{(2)} = S/2$ and the expression for $\delta^2 S$ simplifies to,

$$\begin{aligned}\delta^2 S = \frac{1}{2}[(\frac{\partial^2 S}{\partial U^2})(\delta U^{(1)})^2 + (\frac{\partial^2 S}{\partial V^2})(\delta V^{(1)})^2 + \sum_{\alpha=1}^{c}\sum_{\beta=1}^{c}(\frac{\partial^2 S}{\partial N_\alpha^{(i)}\partial N_\beta^{(i)}})\delta N_\alpha^{(1)}\delta N_\beta^{(1)} \\ +2(\frac{\partial^2 S}{\partial U\partial V})\delta U^{(1)}\delta V^{(1)} + 2\sum_{\alpha=1}^{c}(\frac{\partial^2 S}{\partial U\partial N_\alpha^{(i)}})\delta U\delta N_\alpha^{(1)} + 2\sum_{\alpha=1}^{c}(\frac{\partial^2 S}{\partial V\partial N_\alpha^{(i)}})\delta V^{(1)}\delta N_\alpha^{(1)}]\end{aligned} \tag{24.22}$$

The entropy maximum conditions requires $\delta^2 S \leq 0$ for arbitrary fluctuations $\delta U^{(i)}$, $\delta V^{(i)}$, and $\delta N_\alpha^{(i)}$. If $\delta^2 S = 0$ for some fluctuations, then one needs to examine even higher order variations to explore the stability. We restrict ourselves only to the condition $\delta^2 S \leq 0$, however. In the specific case, $\delta U^{(i)} \neq 0$, $\delta V^{(i)} = 0$, and $\delta N_\alpha^{(i)} = 0$, one need to have $\partial^2 S/\partial U^2 < 0$ satisfied. However,

$$\frac{\partial S}{\partial U} = \frac{1}{T}; \quad \frac{\partial^2 S}{\partial U^2} = -\frac{1}{T^2}\left(\frac{\partial T}{\partial U}\right)\Big|_{T,N}, \tag{24.23}$$

and $C_V = (\partial U/\partial T)|_{N,V}$. Thus, according to the second law of thermodynamics, stable thermal equilibrium requires,

$$C_V > 0. \tag{24.24}$$

Similar analysis of the form given by Eq. (24.22) can be performed under the conditions, $\delta U^{(i)} = 0$, $\delta V^{(i)} \neq 0$, and $\delta N_\alpha^{(i)} = 0$. Then, because $-\partial S/\partial V = P/T$, the second law implies the condition of stable mechanical equilibrium,

$$\left[\frac{\partial(P/T)}{\partial V}\right]\Big|_{U,N} > 0. \tag{24.25}$$

This relation exhibits certain practical inconvenience. Specifically, in the laboratory it is more common to work with the variables T, V, N_α rather than with U, V, N_α. For this reason the conditions of stable equilibrium should be derived using the Helmholtz free energy that attains minimum for a closed isothermal system at temperature T at

equilibrium. In the case of a system with setup like in Fig. 24.1, one needs to explore the condition,.

$$\Delta A = \sum_{i=1,2} [A^{(i)}(U^{(i)} + \delta U^{(i)}, V^{(i)} + \delta V^{(i)}, \mathbf{N}^{(i)} + \delta \mathbf{N}^{(i)}) - A^{(i)}(U^{(i)}, V^{(i)}, \mathbf{N}^{(i)}] \geq 0, \quad (24.26)$$

rather than Eq. (24.10) for entropy. Here, it is important to note, that the condition of thermal equilibrium cannot be derived from the minimum of Helmholtz free energy version of the second law of thermodynamics. This happens because one field variable (the temperature T) is constrained in this version in contrast to $S(U, V, N)$ formulation. Thus, the first variation of $A = A(V, N, T)$ merely recovers the equilibrium conditions in the form $P^{(1)} = P^{(2)}$ and $\mu_\alpha^{(1)} = \mu_\alpha^{(2)}$.

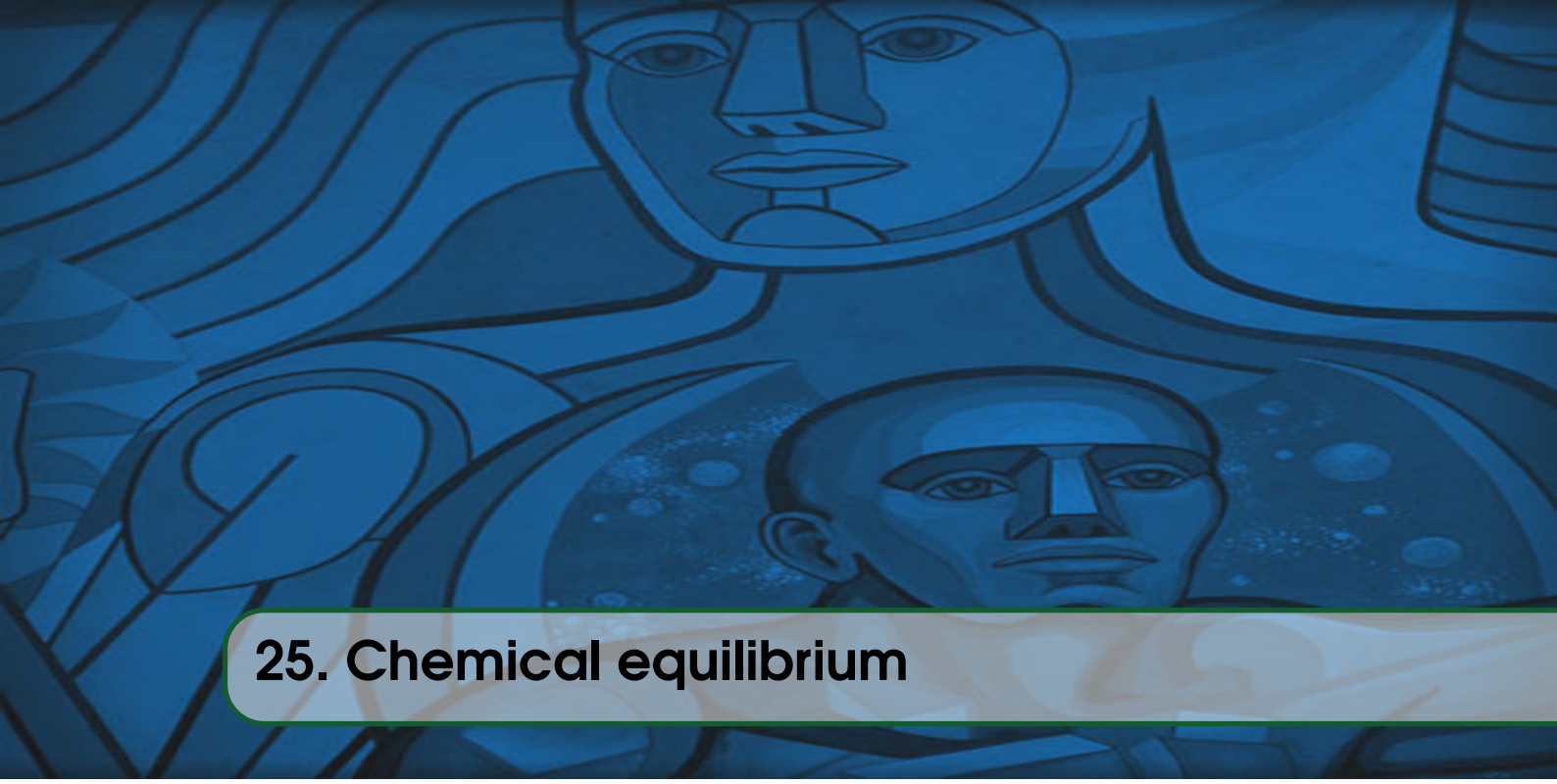
Further, the condition $\Delta A > 0$ for a homogeneous bulk fluid phase (the setup with 1 and 2 considered as identical phases) is assured from the second variation,

$$\delta^2 A = \frac{1}{2}[\frac{\partial^2 A}{\partial V^2})(\delta V^{(1)})^2 + 2\sum_{\alpha=1}^{c}(\frac{\partial^2 A}{\partial V \partial N_\alpha^{(i)}})\delta V^{(1)}\delta N_\alpha^{(1)} + \sum_{\alpha=1}^{c}\sum_{\beta=1}^{c}(\frac{\partial^2 A}{\partial N_\alpha^{(i)}\partial N_\beta^{(i)}})\delta N_\alpha^{(1)}\delta N_\beta^{(1)}] > 0, \quad (24.27)$$

for arbitrary $\delta V^{(1)}$ and $\delta N_\alpha^{(i)}$. Then, the special case $\delta V^{(1)} \neq 0$ and $\delta N_\alpha^{(i)} = 0$ implies that, $\frac{\partial^2 A}{\partial V^2} = -\frac{\partial P}{\partial V} > 0$ and consequently,

$$\left(\frac{\partial P}{\partial V}\right)\Big|_{N,T} < 0. \quad (24.28)$$

It is a more common statement of stable mechanical equilibrium rather than Eq. (24.25). We have used already Eq. (24.28) in the previous chapter while discussing the phase behavior of a fluid in terms of vdW equation of state.



25. Chemical equilibrium

25.1 Introductory remarks

Statistical mechanics of chemical equilibria and of chemical reactions is one of the most important problems at the interface of physics and chemistry due to widespread applications permeating various sciences, e.g. in physical chemistry of inorganic and organic matter and in materials design. Moreover, several research lines within the biological and medicinal chemistry require fundamental knowledge of mechanisms and adequate description of thermodynamics of chemical reactions. The problem in question is interesting and methodologically difficult. Chemical reactions occur in gases, liquids, solids and plasma. Hence, it is necessary to have efficient tools for the description of matter in each of these aggregation states of matter.

For pedagogical reasons, we would like to consider solely some basic issues and focus on the most simple features of methodology concerning most primitive modeling of phenomena. Moreover, our description of the problem is restricted to the chemical association of species in fluids within classical mechanics tools and Boltzmann statistics. One may consider this lecture as an application of the results obtained by using canonical and grand canonical ensemble techniques.

25.2 Statement of the problem

Let us consider a binary system of particles of species A and B with N_A and N_B number of particles, respectively. The external conditions for the system determine its volume V and temperature T. We do not specify the mechanism of chemical reaction occurring in the system, just assume that a "heterogeneous dimerization" reaction of the type,

$$A + B \rightleftharpoons AB, \tag{25.1}$$

occurred and a certain number of molecules (dimers, or products of reaction) AB has been formed due to chemical association of particles or reactants A and B. Note, that AA and/or BB are not formed under our assumption. So, the problem in question is to calculate the number of molecules AB in the system that reached equilibrium at T. In addition, it is necessary to describe thermodynamic properties of this equilibrium system. It is convenient to resort to the grand canonical ensemble defined by μ_A, μ_B, V and T, where μ_A and μ_B are the chemical potentials of species (reactants) A and B.

In close similarity to the derivation of grand canonical ensemble (chapter 11), we introduce the notation ε_i^a, ε_i^b and ε_i^{ab} for the energy levels of a single particle of species A and B and of a dimer (product of reaction) AB. In addition, the notation n_i^a, n_i^b and n_i^{ab} is introduced to describe the occupation numbers of the single-particle energy levels. The A and B type particles can exist as free monomers as well as bounded into dimers AB. Thus, the auxiliary numbers,

$$N_1 = \sum_i n_i^a; \qquad N_2 = \sum_i n_i^b, \tag{25.2}$$

and

$$N_{12} = \sum_i n_i^{ab}, \tag{25.3}$$

correspond to free A and B monomers as well as to AB dimers. With this notation, we can write down the equations,

$$N_A = N_1 + N_{12}; \qquad N_B = N_2 + N_{12}. \tag{25.4}$$

The energy of the system, U, is given by,

$$U = \sum_i \varepsilon_i^a n_i^a + \sum_i \varepsilon_i^b n_i^b + \sum_i \varepsilon_i^{ab} n_i^{ab}. \tag{25.5}$$

It is worth noting, that the last equation is valid if one applies the same zero energy for each term, i.e. the energy levels of dimers should be counted by resting the energy of formation of a dimer of A and B monomers in the ground state.

The canonical partition function for the mixture of monomers and dimers as noninteraction species that obey Boltzmann statistics is,

$$Q = Q(N_1)Q(N_2)Q(N_{12}) = \frac{(\zeta_A)^{N_1}}{N_1!}\frac{(\zeta_B)^{N_2}}{N_2!}\frac{(\zeta_{AB})^{N_{12}}}{N_{12}!}. \tag{25.6}$$

However, as we know from thermodynamics and the theory of ensembles, chemical equilibrium is regulated by the chemical potential. Thus, we proceed to the grand canonical partition function for the system of A and B species that are provided from each of independent respective reservoires (these species next react in the system under consideration),

$$\Xi(\mu_A, \mu_B, V, T) = \sum_{N_A=0}^{\infty}\sum_{N_B=0}^{\infty} Q(N_A, N_B, V, T)\exp(\beta\mu_A N_A)\exp(\beta\mu_B N_B), \tag{25.7}$$

where, $\beta = 1/k_B T$. Now, we substitute Eqs.(25.6) and (25.4) into this expression and collecting the series, obtain,

$$\ln(\Xi) = \zeta_A \exp(\beta\mu_A) + \zeta_B \exp(\beta\mu_B) + \zeta_{AB}\exp(\beta(\mu_A + \mu_B)). \tag{25.8}$$

Note that the information about the products of reaction is hidden in the partition function of the product molecule, ζ_{AB}, and in the chemical potentials that determine the average numbers of particles of species A and B, $< N_A >$ and $< N_B >$. To find them, we apply common relations from the grand canonical partition function in terms of activity, $\lambda_i = \exp(\beta\mu_i)$,

$$< N_i >= \lambda_i \left(\frac{\partial ln\Xi}{\partial\lambda_i}\right), \tag{25.9}$$

with $i = A, B$. Then, it follows,

$$< N_A >= \lambda_A\zeta_A + \lambda_A\lambda_B\zeta_{AB}, \tag{25.10}$$

$$< N_B >= \lambda_B \zeta_B + \lambda_B \lambda_A \zeta_{AB}. \tag{25.11}$$

By comparison of these expressions with Eq.(25.4), we have,

$$< N_1 >= \lambda_A \zeta_A; \qquad < N_2 >= \lambda_B \zeta_B; \qquad < N_{12} >= \lambda_A \lambda_B \zeta_{AB}, \tag{25.12}$$

that yields important result,

$$\frac{< N_{12} >}{< N_1 >< N_2 >} = \frac{\zeta_{AB}}{\zeta_A \zeta_B}. \tag{25.13}$$

The left hand side of this equation expresses the relation between the density of dimers formed in reaction with respect to the product of mononer densities of reagents. It depends on volume, temperature and chemical potentials of species A and B. The r.h.s. is given in terms of these corresponding partition functions.

25.3 Thermodynamic consequences

One should have in mind that laboratory experiment is commonly performed at constant pressure. Therefore, it is of interest to formulate the equation of state of the reacting mixture. One can obtain pressure from the partition function in the gran canonical ensemble, as follows,

$$P = k_B T \left(\frac{\partial \ln(\Xi)}{\partial V} \right) \Big|_{\mu_A, \mu_B, T}, \tag{25.14}$$

or

$$PV = k_B T \ln(\Xi). \tag{25.15}$$

Both lead to the same result,

$$PV = (< N_1 > + < N_2 > + < N_{12} >) k_B T. \tag{25.16}$$

This result has form of the equation of state of a mixture of ideal gases containing however the average values of density of reagents and of the products of reaction. It is worth to complete insights into the thermodynamic properties of an equilibrium state of reacting mixture by examining the free energy.

First, take view of the system in question as a homogeneous macroscopic phase comprising particles of two components, i.e. of A and B species. Then the Gibbs free energy is given by,

$$G = \mu_A < N_A > + \mu_B < N_B > . \tag{25.17}$$

On the other hand, one can express Gibbs free energy of the same system in terms of reacting species and products,

$$G = \mu_A < N_1 > + \mu_B < N_2 > + \mu_{AB} < N_{12} >, \tag{25.18}$$

where μ_{AB} is just a formal notation of the fictitious chemical potential for the products of reaction. Let us make use of the first two equations (25.12), Eqs.(25.12)a and (25.12)b, and sum them. Then,

$$\beta(\mu_A + \mu_B) = \ln(< N_1 >< N_2 >) - \ln(\zeta_A \zeta_B). \tag{25.19}$$

However, the r.h.s. of this equation can be further transformed by using Eq. (25.13) to yield,

$$\beta(\mu_A + \mu_B) = \ln(< N_{12} >) - \ln(\zeta_{AB}). \tag{25.20}$$

This expression can be compared with Eq. (25.12)c expressed in terms of the fictitious chemical potential μ_{AB}, $< N_{12} >= \exp(\beta\mu_{AB})\zeta_{AB}$. Then, we observe that,

$$\mu_{AB} = \mu_A + \mu_B. \tag{25.21}$$

This result shows that the chemical potential of the products equals to the sum of the chemical potentials of reacting species and illustrates conservation of matter assumed in the Eqs.(25.4)a and (25.4)b, confirming, in addition, equivalency of Eqs. (25.17) and (25.18). Thermodynamic results for the Gibbs free energy, Eq. (25.18), and for the equation of state, Eq. (25.16), can be used to obtain Helmholtz free energy of the canonical ensemble, A,

$$A = G - PV = -kT \ln \left(\frac{\zeta_A^{N_1}}{N_1!} \frac{\zeta_B^{N_2}}{N_2!} \frac{\zeta_{AB}^{N_{12}}}{N_{12}!} \right). \tag{25.22}$$

Note, that each average number $< N_i >$ was substituted by constant number N_i, to emphasize that the result corresponds to canonical ensemble conditions. Moreover, now we can rewrite Eq. (25.13) as follows,

$$\frac{N_{12}}{N_1 N_2} = \frac{\zeta_{AB}}{\zeta_A \zeta_B} = K(T,V). \tag{25.23}$$

The function $K(T,V)$ is called the equilibrium rate constant, and the Eq. (25.23) is the so-called mass action law (MAL). Thus, a quantitative description of the chemical reaction for a gaseous mixture implies calculation of the partition functions of reacting species and of products at a given temperature. The densities of reacting species should be known, evidently. It is important to note that one should assume a kind of products resulting from the reaction. Generalization of the procedure to a more complex reaction of the type,

$$\alpha A + \gamma B + .. \rightleftharpoons \eta X + \chi Y + ..., \tag{25.24}$$

leads to the equilibrium rate constant that require calculations of the partition functions of products and reacting species in close similarity to the described in the derivation above,

$$K(T,V) = \frac{(\zeta_X)^\eta}{(\zeta_A)^\alpha} \frac{(\zeta_Y)^\chi}{(\zeta_B)^\gamma} ... \tag{25.25}$$

Again, the type of products of the reaction should be known. To summarize the developments above, we would like to mention a few important issues. In the course of theoretical explanation, we restricted to gaseous phase and used Boltzmann statistics. Most important is that the Hamiltonian of the problem does not contain interaction between particles. In other words the only possible mechanism of the formation of products between "reactive" particles is via their collisions. A more general consideration of chemical association effects of species in e.g. dense gases and/or liquids require formulation of the potential energy term of the Hamiltonian. At this level of modelling, even if one restricts to classical statistical mechanics and Boltzmann statistics, the problem become much more difficult. For pedagogical reasons we would like to state the problem in a simple manner, such that the reader solely would get flavor of necessary theoretical developments. The most valuable contributions into this area of research were elaborated a few decades ago by M. S. Wertheim [37, 38, 39, 40]. Here, in the following subsection we would like to consider some of the modeling issues from a less sophisticated approach by G. Stell and P. Cummings [41].

Before that, it is worth mentioning that strictly speaking the interaction potentials to yield association between particles of species A and B must result from quantum mechanical considerations. Nevertheless, using effective pair potentials as input to a classical statistical mechanical procedure has a long history. One of interesting and

important examples is the so-called central force model for liquid water by Lemberg and Stillinger. In this model, water is considered as a mixture of charged hydrogen, H^q, and oxygen, O^{-2q}, atoms interacting via pairwise potentials with $q = 0.32983e_0$ (e_0 is the elementary charge). The interaction between positively charged hydrogen atom and negatively charged oxygen atom (both with "effective" charges) is Coulomb-like at large separations and is rapidly varying function with specific shape at short range. This interaction, together with the repulsive potential of peculiar shape between hydrogen atoms, imposes correct geometry and stoichiometry of the water molecules (see Fig. 25.1).

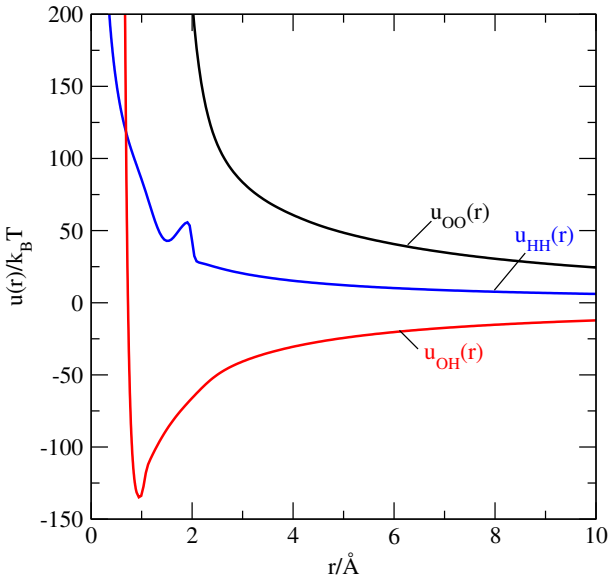


Fig. 25.1: The central force model of water, $u(r)/k_BT$, at $T = 298.15$ K for the intramolecular forces oxygen-oxygen (OO), hydrogen-hydrogen (HH), and oxygen-hydrogen (OH) [42].

Statistical mechanical treatment of the model leads to adequate description of the network of hydrogen bonds in liquid water at room temperature and reasonable values of thermodynamic properties.

25.4 On the chemical association in non-ideal fluids

As mentioned slightly above, we would like to elaborate here the same example, as in Sec. 25.2, of the simple chemical reaction, $A + B \rightleftharpoons AB$, and focus in the problem to get the equilibrium rate constant, written similarly to Eq. (25.23),

$$K(T,V) = \frac{[AB]}{[A][B]}, \tag{25.26}$$

where the notation $[X]$ corresponds to the concentration of species X particles. Thus, we would like to consider heterogeneous dimerization reaction. The potential energy of the system consists of the pair interaction potentials (similarly to Eq. (6.13) of chapter 6) between particles of all species, $u_{AA}(r)$, $u_{BB}(r)$ and $u_{AB}(r)$. We assume equimolar mixture of A and B. Formally speaking, one can calculate the pair distribution functions, $g_{AA}(r)$, $g_{BB}(r)$ and $g_{AB}(r)$, by using integral equations method (see chapter 23) that describe the microscopic structure and thermodynamics of the model. The probabilistic interpretation of the $g_{AB}(r)$ distribution function permits to obtain the average number e.g. of B type particles whose distance from a chosen A particle is in the interval between l_1 and l_2 (introduce notation $n^B_{l_1 l_2}$) as follows,

$$n^B_{l_1 l_2} = \rho_{B0} \int_{l_1<r<l_2} g_{AB}(r) d\mathbf{r}, \tag{25.27}$$

where $\rho_{B0} = \rho_{A0}$, are the total densities of species B and A particles. Then, it is natural to define a pair of A and B to be "chemically" associated into a dimer (or diatomic) molecule AB by using geometric criterion,

$$\rho_{AB} = \rho_{A0}\rho_{B0} \int_{R_d<r<R_u} g_{AB}(r) d\mathbf{r}, \tag{25.28}$$

where the existence of a molecule AB with the intramolecular parameter, r_{AB}, is identified by the interval of distances $R_d < r < R_u$ (down and up). This equation provides definition of the quantity $[AB]$ in Eq. (25.26). The quantities $[A]$ and $[B]$ are identified with the densities of non-associated (free) A and B particles. In essence, the equilibrium rate constant is related to the pair distribution function.

As we have mentioned in the previous section the kind of chemical reaction to deal with determines the partition function of the product. Here, this statement must be reformulated as follows. One should design the form of interaction potential that fit or say satisfy conditions of the formation of definite type of products. In other words, the Hamiltonian model must ensure steric saturation for the reaction in question, that is the formation of diatomic molecules but not the formation of n-mers for $n \geq 3$ (e.g. ABA, AAB, BBA, etc.). The prototype model for dimerization, applied by Cumming and Stell [41], is,

$$u_{AA}(r) = u_{BB}(r) = \begin{cases} \infty, & \text{if } r < \sigma, \\ 0, & \text{if } r \geq \sigma, \end{cases} \tag{25.29}$$

and

$$u_{AB}(r)/k_B = \begin{cases} \varepsilon_1, & 0 < r < L - w/2, \\ -\varepsilon_2, & L - w/2 < r < L + w/2, \\ \varepsilon_1, & L + w/2 < r < \sigma, \\ 0, & r > \sigma, \end{cases} \tag{25.30}$$

where k_B is Boltzmann's constant. If R_u is chosen as $L + w/2$ and R_d to be in the interval $[0, L + w/2)$, the definition for the density of dimers according to Eq. (25.28) is unambiguous through the prevention of polymerization beyond the dimer level. Once a type A and type B particles are separated by less than $L + w/2 \leq \sigma/2$, no other type A or type B particles may be bound to the dimer, according to the geometry of interaction. This mechanism of excluding species with three and more monomer can be referred to as steric saturation.

The choice $R_d = L - w/2$ is quite natural, since it corresponds to an energetic definition of associated atoms. The two atoms are defined to be associated if their two-body energy is large and negative. One can consult a more references to find different versions of criteria to describe e.g. hydrogen bonding in water and aqueous solutions.

Unfortunately, the model given by Eqs. (25.29) and (25.30) is not solvable analytically for technical reasons. Therefore, to get analytical solution, Cummings and Stell had transformed the model into infinitely deep ($\varepsilon_2 \to \infty$), infinitesimally wide ($w \to 0$) attractive well, keeping the second virial coefficient (cf. Eq. (21.39) from chapter 21) of the original and modified model equal. The pair potential model used for the AB interaction is depicted in Fig. 25.2.

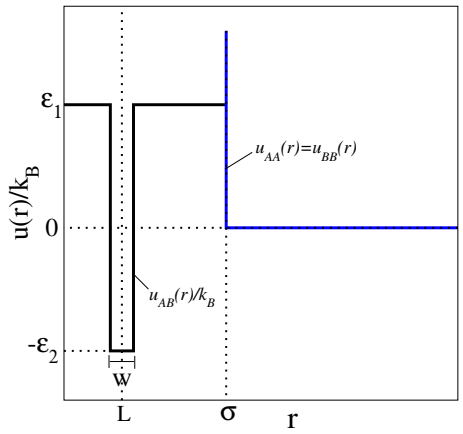


Fig. 25.2: Prototype model for dimerization [41]. The blue line corresponds to the AA and BB interactions (hard sphere potential with diameter σ) and the black line corresponds to the AB interaction.

Next, the modified model has been solved by using integral equations methodology of the theory of liquids to get $g_{AB}(r)$ and elaborate the equilibrium rate constant as a function of density and temperature according to Eq. (25.28). The level of mathematics necessary to follow the solution in details does not permit us to provide a more extended explanation in this text. Interested students can consult Refs. [41, 42, 43, 44] and the references therein.

26. Appendix

26.1 Appendix A: Thermodynamic relations

In these notes, we have used the following thermodynamic relation,

$$C_p - C_v = \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right]\left(\frac{\partial V}{\partial T}\right)_P. \tag{26.1}$$

It is an interesting equation deserving analysis in more detail. We provide a set of comments on it below.

Let us begin with with some auxiliary issues. Namely, formal relation determining equation of state of a fluid is, $P = P(V,T)$. Then,

$$dP = \left(\frac{\partial P}{\partial V}\right)_T dV + \left(\frac{\partial P}{\partial T}\right)_V dT. \tag{26.2}$$

If one works at constant pressure ($dP = 0$), then,

$$\left(\frac{\partial P}{\partial V}\right)_T dV = -\left(\frac{\partial P}{\partial T}\right)_V dT, \tag{26.3}$$

or

$$\left(\frac{dV}{dT}\right)_P = -\left(\frac{\partial P}{\partial T}\right)_V \Big/ \left(\frac{\partial P}{\partial V}\right)_T. \tag{26.4}$$

In other words, we have,

$$\left(\frac{\partial P}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P \left(\frac{\partial T}{\partial P}\right)_V = -1. \tag{26.5}$$

Now, we would like to refer to experimentally accessible or say measurable quantities. Namely, $\alpha = (1/V)(\partial V/\partial T)_P$ is called the temperature coefficient of volume expansion, whereas $\beta = -(1/V)(\partial V/\partial P)_T$ is the isothermal compressibility.

From the first law of thermodynamics, $dQ = dU + dW$ ($dW = PdV$, $U = U(T,V)$), we have,

$$dQ = dU + PdV = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV + PdV, \tag{26.6}$$

or

$$dQ = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] dV. \tag{26.7}$$

Recall that the heat capacity at constant volume is: $C_V = (dQ/dT)_V = (dU/dT)_V$. Now, we would like to make transformation to the independent variables, P and T, using $V = V(T,P)$, that correspond to the conditions of laboratory experiment. Hence,

$$dQ = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] \left(\frac{\partial V}{\partial T}\right)_P dT. \tag{26.8}$$

This relation yields the heat capacity at constant pressure,

$$C_P = \left(\frac{dQ}{dT}\right)_P; \qquad C_P - C_V = \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] \left(\frac{\partial V}{\partial T}\right)_P, \tag{26.9}$$

How we can use this relation and verify it by measurements? From the second law of thermodynamics we use the concept of entropy, $dQ = TdS$. Combining what we had already above from the first law i.e. from Eq. (26.7),

$$\left(\frac{\partial S}{\partial T}\right) = \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V, \tag{26.10}$$

and

$$\left(\frac{\partial S}{\partial V}\right) = \frac{1}{T}\left[\left(\frac{\partial U}{\partial V}\right)_T + P\right]. \tag{26.11}$$

However, $\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T}$. Explicitly, we obtain,

$$\frac{\partial}{\partial V}\left(\frac{\partial S}{\partial T}\right) = \frac{1}{T}\frac{\partial^2 U}{\partial V \partial T}, \tag{26.12}$$

and

$$\frac{\partial}{\partial T}\left(\frac{\partial S}{\partial V}\right) = -\frac{1}{T^2}\left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] + \frac{1}{T}\left[\frac{\partial^2 U}{\partial T \partial V} + \left(\frac{\partial P}{\partial T}\right)_V\right]. \tag{26.13}$$

Then, it follows,

$$-\left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] + T\left(\frac{\partial P}{\partial T}\right)_V = 0. \tag{26.14}$$

Therefore, we can eliminate the square bracket in Eq. (26.9) in favor of the derivative of pressure by temperature,

$$C_P - C_V = T\left(\frac{\partial P}{\partial T}\right)_V \left(\frac{\partial V}{\partial T}\right)_P = -T\left[\left(\frac{\partial V}{\partial T}\right)_P\right]^2 \left(\frac{\partial P}{\partial V}\right)_T = VT\alpha^2/\beta, \tag{26.15}$$

where the result of chain Eq. 26.10 has been used. In conclusion, the l.h.s. of the equation above is expressed only in terms of experimentally measurable coefficients.

26.2 Appendix B: Elements of the probability theory

In many phenomena, as it follows from practice or real life, one can observe not entirely definite results upon repetition of the same action (or say performing the same experiment) under the same conditions. In other words, one is speaking about the experiments yielding random results. Let call them as random experiments and assume that they can be repeated, at least in principle, unlimited number of times. Mathematical description of such random experiments of phenomena is the subject of the probability theory.

Any kind of results of such random experiment observed either visually or by using certain measurement device can be denominated as a random event. A few examples can be given, such as: A) observing head or tail upon throwing a coin, B) winning the lottery, C) amount of goods sold in a shop per day. Intuitively clear that the event A is more possible than the event B, for example.

One would like to introduce mathematical or number characteristics for the possibility of occurrence of random event in a certain experiment. Possibly the simplest way is to perform N times the same experiment and count the number $N(A)$ when the event A happens. Then, the probability of the event A, called $P(A)$ is the number related to the frequency of this event in the following form,

$$P(A) = \lim_{N\to\infty} \frac{N(a)}{N}. \tag{26.16}$$

Evidently, that the definition implies, $0 \leq P(A) \leq 1$. If the event A occurs always, then $P(A) = 1$, in contrast $P(A) = 0$, if A never occurs.

Additional elements used below are as follows. One can define the space of elementary events, Ω. Those are definite distinct results, that mutually exclude each other. As an example, consider an experiment of throwing a true die once.

Then the elementary events can be $\omega_i = 1,2,3,4,5,6$ describing outputs when the number i appears on the upper surface of the die. Consequently, the space of elementary events in this particular case is $\Omega = \omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6$.

If the space Ω is discrete (finite or can be enumerated), then the random event represents any subset of Ω. Moreover, any random event $A \subset \Omega$ consists of elementary events. Referring to the example of a die above, a subset of elementary events that form event A =(odd number of points) is $A = \omega_1,\omega_3,\omega_5$; or e.g. event B =(the number of points on the edge is greater than three) is $B = \omega_4,\omega_4,\omega_6$.

By such definition the event is identified with certain set, thus one can perform operations, like sum, difference and product, with events as are done with sets of elements. In other words, one can define algebra of events with necessary properties. To explain this, let us explore the simplest exercise. Namely, a single experiment of throwing the coin is done. The task is to construct the space of elementary events and to construct algebra in it. Possible results of the experiment determine two elementary events: H (head) and T (tail). The space of elementary events is: $\Omega = \{H,T\}$. The space Ω is finite, therefore, as algebra $\mathbf{A}$, we can choose a set of all possible subsets from Ω. Thus, $\mathbf{A} = (\emptyset,H,T,\{H,T\})$, and satisfies all the properties of algebra.

Three elements, Ω, $\mathbf{A}$ and $P(A)$ - a function defined on the algebra $\mathbf{A}$ and describing the probability of the event A all together are called the probabilistic space and represent a mathematical model of random experiment.

Several experimental situations, like throwing a coin or a die, or some others, are well described by simple models. Really, if a set of elementary events is finite (and in fact not too big) like in these mentioned cases, let us denote the probabilities of the elementary events, ω_i, $i = 1,2,3,4,..,n$, $\omega_i \in \Omega$ by non-negative numbers $p(\omega_i)$, such that,

$$\sum_{i=1}^{n} p(\omega_i) = 1, \tag{26.17}$$

then the probability $P(A)$ of the event A ($A \subset \Omega$), can be determined from the equation,

$$P(A) = \sum_{\omega_i \in A} p(\omega_i), \tag{26.18}$$

where ω_i are the elements of A. Of particular interest is the case when the probabilities of all elementary events are equal, i.e. $p(\omega_i) = p$, for all i, $i = 1,2,3,4,..,n$. Then, $p = 1/n$. More generally, the equation above can be rewritten, as,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{k}{n}, \tag{26.19}$$

where $|A| = k$ is the number of elementary events that for A, and $|\Omega| = n$ the entire number of elementary events. This expression corresponds to the so-called classical definition of probability.

If, one needs to calculate explicitly, the numbers k and n to find $P(A)$, the combinatorial formulas are used. Recall that the number of m-combinations of an n-set, C_n^m, is related to the number of m-permutations of n by,

$$C_n^m = \frac{n!}{m!(n-m)!}. \tag{26.20}$$

As an example consider the following exercise. There are 20 identical milk bottles and 5 of them contain bad or spoiled (or overtimed) product. To make quality test, 3 bottles have been randomly chosen. What is the probability that only the good bottles will be between these three? As an elementary events in such an experiment, we have the choice of any triples of 20. All these events are equiprobable and their number is $|\Omega| = C_{20}^3$. The number of elementary events forming the event A equals to C_{15}^3. Then,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{C_{15}^3}{C_{20}^3} = \frac{15!/3!12!}{20!/3!17!} = 91/228. \tag{26.21}$$

The classic definition of probability cannot be used to the experiment with infinite number of equiprobable results in Ω. In such a case, one needs to resort to the geometric definition of probability. Assume that Ω is the restricted region on a) number axis, b) on plane and c) in three-dimensional space. Assume that Ω has a) length, b) area and c) volume. A system $\mathbf{A}$ of subsets from Ω form an adequate algebra to define the probability. Then, by definition, the event A, has the probability,

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}, \tag{26.22}$$

where $\mu(A)$ is the length or area or volume of a subset A. This is a geometric definition. For illustration make the following exercise. Let throw a point into the circle with radius R. How to find the probability that the point will fall into a square, centered inside a circle and having a diagonal with the length $2R$. Show that $P(A) = 2/\pi$.

In several places of the previous discussion we referred to the results of a single experiment to apply certain mathematical constructions. On the other hand, a set of experiments is necessary to perform with the sake of using concepts of the probability theory. What we have in mind is that if a series of experiments is performed, the results of each of them appear to be independent, or in other words the result of a given experiment or trial is independent on the result of the previous experiment. The simplest class of repeated trials is a series of independent trials with solely two possible outputs, e.g. success of failure. As a success we think off the occurrence of certain event A, and by failure - that the event A does not occur in a single trial. The probabilities of success $P(A) = p$ and of failure $P(A) = q = 1 - p$ are constant along the entire series of tests. Such a series is called as the scheme of independent trials of Bernoulli. The formula of Bernoulli

determines the probability $P_n(m)$ of observing the event A m times while performing n independent trials,

$$P_n(m) = C_n^m p^m q^{n-m}, \tag{26.23}$$

where C_n^m is the combinatorial multiplier given above.

If n and m are sufficiently big, then it is very difficult to use the formula of Bernoulli in practical calculations. In practice, several approximations are used dependent on the values of p,q,n,m. Namely, if p is small ($p < 0.1$) and for big n, such that $np < 20$, the Poisson expression is conveniently used,

$$P_n^m \approx \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda = np. \tag{26.24}$$

On the other hand, if $n > 100$, p and q are not small and if $npq > 20$, the approximate formula of Moivre - Laplace is applied,

$$P_n^m \approx \frac{1}{\sqrt{npq}} \phi(x_m), \quad x_m = \frac{m - np}{\sqrt{npq}}, \tag{26.25}$$

where the function $\phi(x)$ is,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{26.26}$$

It is worth mentioning the properties of the function $\phi(x)$. It is even, $\phi(x) = \phi(-x)$ and takes maximum value at $x = \frac{1}{\sqrt{2\pi}}$. The function can be plotted easily.

In many applications it is necessary to have the probability, that the event will occur not less than m_1 times and not more than m_2 times, i.e. to calculate $P_n(m_1 < m < m_2)$. The result of this type of calculations follows from the integral theorem of Moivre - Laplace. Namely, if p is constant ($0 < p < 1$), then at $n \to \infty$,

$$P_n\left(a \leq \frac{m - np}{\sqrt{npq}} \leq b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx. \tag{26.27}$$

For the sake of better understanding, let us comment on the properties of the Laplace function used just above,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}x^2} dx. \tag{26.28}$$

It is easy to observe that, this function is even, $\Phi(x) = \Phi(-x)$, $\Phi(0) = 0$, the function $\Phi(x)$ is monotonously growing and $\lim_{x\to\infty} \Phi(x) = 0.5$. Consequently,

$$P_n\left(a \leq \frac{m - np}{\sqrt{npq}} \leq b\right) \to \Phi(b) - \Phi(a). \tag{26.29}$$

In several places and exercises above we have used issues just reformulated and stated more explicitly now. A random variable $\xi(\omega)$ is the number function of the elementary event $\omega \in \Omega$. It is in fact an extension of the notion of number function, $y = f(x)$, of the real argument x in calculus. Because $\xi(\omega)$ is the number function it has correspondence with a set of numbers of real number axis R. These numbers describe a set of possible values of the random variable. More important is that to describe the random variable completely, it is insufficient to give solely a set of its possible values. Namely, in addition, it is necessary to provide the probabilities of all possible values. Any kind of a rule or

law, that establishes relation between possible values of random variable is called as the law of distribution (or just distribution) of random variable. Dependent on the intrinsic nature of the set of possible values one can distinguish random variable of discrete type and continuous random variable. For example, in close similarity to our discussion above, one can say that the random discrete variable $\xi(\omega)$ is distributed according to the Poisson law, if its distribution has the form

$\xi(\omega)$	0	1	2	3	...
P_m	P_0	P_1	P_2	P_3	...

where $P_m = \frac{\lambda^m}{m!} e^{-\lambda}$ $\lambda = np$. It is easy to show that $\sum_{k=0}^{\infty} P_k = 1$.

Proceed now to the continuous random variable. A random variable $\xi(\omega)$ is called random continuous variable if a set of its values represents certain interval (a,b) (finite or infinite) of the number axis R. Because possible values of random continuous variable continuously fill the interval (a,b), the construction of the distribution should differ from the case of a discrete variable. The method in the present case is in the construction of a distribution function.

Definition: The function $F_{\xi}(x)$ of the real variable x, $-\infty < x < \infty$, defined by the expression,

$$F_{\xi}(x) = P(\xi < x), \tag{26.30}$$

is called the distribution function of the random variable $\xi(\omega)$.

Thus, $F_{\xi}(x)$ is the probability that $\xi(\omega)$ takes value less than x (or takes value in the interval $-\infty < x$). With this definition, one can find the probability of any kind of event of the type $x_1 \leq \xi(\omega) < x_2$ or in other words from the interval $[x_1, x_2)$. To do that, let us use the following relation between events,

$$\xi(\omega) < x_2 = \xi(\omega) < x_1 + (x_1 \leq \xi(\omega) < x_2). \tag{26.31}$$

Then, it follows that,

$$F_{\xi}(x_2) = F_{\xi}(x_1) + P(x_1 \leq \xi(\omega) < x_2), \tag{26.32}$$

and consequently,

$$P(x_1 \leq \xi(\omega) < x_2) = F_{\xi}(x_2) - F_{\xi}(x_1). \tag{26.33}$$

It is easy to note that: $0 \leq F_{\xi}(x) \leq 1$ for $-\infty < x < \infty$; $F_{\xi}(-\infty) = 0$ and $F_{\xi}(\infty) = 1$. Moreover, for continuous random variable $\xi(\omega)$ its distribution function $F_{\xi}(x)$ is continuous as well.

Final comments of this chapter refer to one more important concept. First, we would like to mention that in the majority of applications most frequently continuous random variables are the *absolutely* continuous random variables. By definition continuous random variable $\xi(\omega)$ is called the *absolutely* continuous random variable if for all values x of the number axis R there exists a function $f_{\xi}(x)$ such that,

$$F_{\xi}(x) = \int_{-\infty}^{x} f_{\xi}(t)dt. \tag{26.34}$$

The function $f_{\xi}(x)$ is called the probability density distribution for the random variable $\xi(\omega)$. The probability density has the following properties: 1) $f_{\xi}(x) \geq 0$, $-\infty < x < \infty$; 2) Normalization condition $\int_{-\infty}^{\infty} f_{\xi}(t)dt = 1$; 3) $F'_{\xi}(x) = f_{\xi}(x)$ in the points of continuity of $f_{\xi}(x)$.

Using the definition, we can rewrite Eq. (26.33) as follows,

$$P(x_1 \leq \xi(\omega) < x_2) = \int_{x_1}^{x_2} f_{\tilde{\xi}}(t)dt. \tag{26.35}$$

The notion of the probability density distribution becomes clearer then. Namely, the integration of the probability density yields probability on the l.h.s. of this equation.

Two exercises are provided as a homework.

Exercise 26.1 Let the probability density $f_{\tilde{\xi}}(x)$ of $\xi(\omega)$ is,

$$f_{\tilde{\xi}}(x) = \begin{cases} asinx, & 0 \leq x \leq \pi \\ 0, & x < 0, x > \pi \end{cases}, \tag{26.36}$$

Find: a) the coefficient a; b) the distribution function $F_{\tilde{\xi}}(x)$ and c) the probability that $\xi(\omega)$ is in the interval $(0, \pi/4)$.

Solution: The probability density function should be normalized! Then,

$$\int_{-\infty}^{\infty} f_{\tilde{\xi}}(x)dx = \int_0^{\pi} asinxdx = 1. \tag{26.37}$$

Then,

$$a = 1/\int_0^{\pi} sinxdx = 1/2, \tag{26.38}$$

$$F_{\tilde{\xi}}(x) = \int_{-\infty}^{x} f_{\tilde{\xi}}(t)dt = (1/2)\int_0^{x} sintdt = (1 - cosx)/2, \tag{26.39}$$

where $0 < x \leq \pi$. Besides, $F_{\tilde{\xi}}(x) = 0$ for $x \leq 0$ and $F_{\tilde{\xi}}(x) = 1$ for $x > \pi$. Result for point c is: $P \approx 0.147$.

■

Exercise 26.2 A random variable is $\xi(\omega)$ is called as homogeneously distributed in the interval $[a, b]$, if its density distribution has the form,

$$f_{\tilde{\xi}}(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & x < a, x > b \end{cases}, \tag{26.40}$$

With this assumption, find the distribution function, $F_{\tilde{\xi}}(x)$, and plot it. Result: $F_{\tilde{\xi}}(x) = 0$ for $x \leq a$ and $F_{\tilde{\xi}}(x) = 1$ for $x > b$, $F_{\tilde{\xi}}(x) = (x - a)/(b - a)$ for $a < x \leq b$.

■

26.3 Appendix C: Mathematical Series and Special Functions

The Taylor expansion of the logarithm is written as,

$$\ln(1 + x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots, \tag{26.41}$$

for $|x| < 1$. The geometric series can be written as,

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \tag{26.42}$$

convergent for $|x| < 1$. The exponential function e^x can be expressed as a series, as follows,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{26.43}$$

where the serie converges for all value of x. The usual trigonometric functions and their inverses have the following series form,

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots, \tag{26.44}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots, \tag{26.45}$$

$$\tan x = \sum_{n=1}^{\infty} \frac{B_{2n}(-4)^n(1-4^n)}{(2n)!} x^{2n-1} = x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots, \tag{26.46}$$

where B_k's are the Bernoulli numbers. In addition to these equations, we have included the following hyperbolic functions,

$$\tanh x = \sum_{n=1}^{\infty} \frac{2^{2n}(2^{2n}-1)B_{2n}x^{2n-1}}{(2n)!} = x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + \cdots, \tag{26.47}$$

vaid for $|x| < \frac{\pi}{2}$. For $\coth x$, the Taylor series expansion is given by,

$$\coth x = \sum_{n=0}^{\infty} \frac{2^{2n}B_{2n}x^{2n-1}}{(2n)!} = \frac{1}{x} + \frac{x}{3} - \frac{x^3}{45} + \frac{2x^5}{945} - \frac{x^7}{4725}\cdots, \tag{26.48}$$

valid for $0 < |x| < \pi$. One of the most important special functions is the Γ function. The gamma function, $\Gamma(z)$, is defined for all complex numbers z except non-positive integers. For every positive integer $z = n$, it satisfies,

$$\Gamma(n) = (n-1)!. \tag{26.49}$$

The gamma function can be defined via a convergent improper integral for complex numbers with positive real part,

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}\,\mathrm{d}t, \qquad \Re(z) > 0\,. \tag{26.50}$$

If the real part of the complex number z is strictly positive ($\Re(z) > 0$), then the integral of Eq. (26.50) converges absolutely, and is known as the Euler integral of the second kind.

Other important special function is the Riemann zeta function, denoted by the letter ζ. This function is defined as,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots, \tag{26.51}$$

for $\mathrm{Re}(s) > 1$. The Riemann zeta function is a function of a complex variable $s = \sigma + it$, where σ and t are real numbers. When $Re(s) = \sigma > 1$, the function can be written as a converging summation or as an integral,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{x^{s-1}}{e^x - 1}\,\mathrm{d}x\,. \tag{26.52}$$

The Riemann zeta function is defined for other complex values via analytic continuation of the function defined for $\sigma > 1$.

26.4 Appendix D: Elements of quantum mechanics

This piece below is from Russian brief version of L. Landau, E. Lifshitz "Quantum mechanics" [10].

In matrix mechanics, observables such as position and momentum are represented by self-adjoint operators. When considering pairs of observables, an important quantity is the commutator. For a pair of operators $\hat{A}$ and $\hat{B}$, one defines their commutator as,

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{26.53}$$

In the case of position and momentum, the commutator is the canonical commutation relation,

$$[\hat{x}, \hat{p}] = i\hbar. \tag{26.54}$$

The physical meaning of the non-commutativity can be understood by considering the effect of the commutator on the coordinate and momentum eigenstates. Let $|\psi\rangle$ be a right eigenstate of the coordinate with a constant eigenvalue x_0. By definition, this means that $\hat{x}|\psi\rangle = x_0|\psi\rangle$.

Applying the commutator to $\psi\rangle$ yields,

$$[\hat{x}, \hat{p}]|\psi\rangle = (\hat{x}\hat{p} - \hat{p}\hat{x})|\psi\rangle = (\hat{x} - x_0\hat{I})\hat{p}|\psi\rangle = i\hbar|\psi\rangle, \tag{26.55}$$

where $\hat{I}$ is the identity operator.

Suppose, for the sake of proof by contradiction, that $|\psi\rangle$ is also a right eigenstate of momentum, with constant eigenvalue p_0. If this were true, then one could write,

$$(\hat{x} - x_0\hat{I})\hat{p}|\psi\rangle = (\hat{x} - x_0\hat{I})p_0|\psi\rangle = (x_0\hat{I} - x_0\hat{I})p_0|\psi\rangle = 0. \tag{26.56}$$

On the other hand, the above canonical commutation relation requires that,

$$[\hat{x}, \hat{p}]|\psi\rangle = i\hbar\psi\rangle \neq 0. \tag{26.57}$$

This implies that no quantum state can simultaneously be both a coordinate and a momentum eigenstate.

When a state is measured, it is projected onto an eigenstate in the basis of the relevant observable. For example, if a particle's position is measured, then the state amounts to a position eigenstate. This means that the state is not a momentum eigenstate, however. Rather it can be represented as a sum of multiple momentum basis eigenstates. In other words, the momentum must be less precise. This precision may be quantified by the standard deviations,

$$\sigma_x = \sqrt{\langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2}, \tag{26.58}$$

and

$$\sigma_p = \sqrt{\langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2}. \tag{26.59}$$

One observes a balance between the respective precisions of the two, quantified by the uncertainty principle,

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \tag{26.60}$$

26.5 Appendix E: linear transformations

First, let us introduce the variables, $\eta_i = (m_i/m)^{1/2}\xi_i$ and $\tilde{a}_{ij} = (m^2/m_i m_j)^{1/2}a_{ij}$, and make the transformation of the initial Hamiltonian defined by Eq. (14.11) and given in terms of variables ξ, into the form,

$$H = \sum_{i=1}^{3N} \frac{1}{2} m\dot{\eta}_i^2 + \frac{1}{2}\sum_{i,j}^{3N} \tilde{a}_{ij}\eta_i\eta_j + U_N^{(0)}. \tag{26.61}$$

The matrix $\mathbf{A}$ with the elements $\tilde{a}_{ij}$ is symmetric in close similarity to the initial matrix a_{ij}. Moreover its elements are real. In consequence, the matrix $\mathbf{A}$ is self-adjoint. Below, one can find some useful notes. Next, we return to the problem in question.

Note 1: A square matrix is called Hermitian if it is self-adjoint. Therefore, a Hermitian matrix $\mathbf{A} = (a_{ij})$ is defined as one for which $\mathbf{A} = \mathbf{A}^H$, where $\mathbf{A}^H$ denotes the conjugate transpose. This is equivalent to the condition $a_{ij} = \bar{a}_{ji}$, where $\bar{z}$ denotes the complex conjugate. As a result of this definition, the diagonal elements a_{ii} of a Hermitian matrix are real numbers (since $a_{ii} = \bar{a}_{ii}$), while other elements may be complex.

Note 2: Eigenvalues are a special set of scalars associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic roots, characteristic values, proper values, or latent roots.

Note 3: The determination of the eigenvalues and eigenvectors of a system is extremely important in physics and engineering, where it is equivalent to matrix diagonalization and arises in such common applications as stability analysis, the physics of rotating bodies, and small oscillations of vibrating systems, to name only a few. Each eigenvalue is paired with a corresponding so-called eigenvector.

Mathematical notes: Let A be a linear transformation represented by a matrix $\mathbf{A}$. If there is a vector $\mathbf{X}$ in $\mathbf{X} \in \mathbf{R}^n \neq 0$ such that,

$$\mathbf{A}\mathbf{X} = \lambda\mathbf{X}, \tag{26.62}$$

for some scalar λ, then λ is called the eigenvalue of $\mathbf{A}$ with corresponding eigenvector $\mathbf{X}$. Let A be a square matrix with eigenvalue λ. Then, the corresponding eigenvectors satisfy the equation,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = 0, \tag{26.63}$$

where $\mathbf{I}$ is the identity matrix. A linear system of equations above has nontrivial solutions if the determinant vanishes, so the solutions of the matrix equation are given by,

$$det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{26.64}$$

This equation is known as characteristic equation of $\mathbf{A}$.

For example, for a 2×2 matrix, the eigenvalues are,

$$\lambda_{1,2} = \frac{1}{2}\left[(a_{11} + a_{22}) \pm \sqrt{(4a_{12}a_{21} + (a_{11} - a_{22})^2)}\right], \tag{26.65}$$

which arises as the solutions of the characteristic equation,

$$x^2 - x(a_{11} + a_{22}) + (a_{11}a_{22} - a_{12}a_{21}) = 0. \tag{26.66}$$

If the eigenvalues are known, then the eigenvectors corresponding to each eigenvalue can be found using Eq. (26.63).

Return now, to the transformations of our interest. In order to perform diagonalization of the term describing the potential energy $\eta_i\eta_j$ in the Hamiltonian let do the following steps. The self adjoint square matrix we deal with, is of the size $3N \times 3N$. So it has $3N$ orthonormal eigenvectors $\mathbf{s}_i$ that can be determined from matrix equation

$$\mathbf{A}\mathbf{s}_i = \lambda_i\mathbf{s}_i, \tag{26.67}$$

where λ_i is the i-th eigenvalue of the matrix $\mathbf{A}$. The eigenvector $\mathbf{s}_i$ is a column of $3N$ elements, denote them $s_{i1}, s_{i2}, ..., s_{i,3N}$. The eigenvectors of the matrix in question are orthonormal, i.e.,

$$\mathbf{s}_i^T \mathbf{s}_j = \sum_{k=1}^{3N} s_{ik} s_{jk} = \delta_{ij}. \tag{26.68}$$

In above, $\mathbf{s}_i^T$ is the row vector constructed by taking transpose of the column vector $\mathbf{s}_i$. Now, let us form a matrix $\mathbf{S}$ from the column eigenvectors $\mathbf{s}_1, ..., \mathbf{s}_{3N}$. This matrix has the following properties,

$$\mathbf{S}^T \mathbf{S} = \mathbf{I} = [\delta_{ij}], \tag{26.69}$$

and,

$$\mathbf{S}^T \mathbf{A} \mathbf{S} = \mathbf{\Lambda} = [\lambda_i \delta_{ij}]. \tag{26.70}$$

These manipulations permit us to make the following steps. First, rewrite the Hamiltonian of Eq. (26.61) using matrix notation,

$$H = \frac{1}{2} m \dot{\boldsymbol{\eta}}^{\boldsymbol{T}} \dot{\boldsymbol{\eta}} + \frac{1}{2} \boldsymbol{\eta}^{\boldsymbol{T}} \boldsymbol{A} \boldsymbol{\eta} + U_N^{(0)}, \tag{26.71}$$

and then apply transformation of coordinates, $\boldsymbol{\eta} = \boldsymbol{S}\boldsymbol{\zeta}$ in the spirit of Eq. (14.13). Then, the Hamiltonian becomes converted into the desired result,

$$H = \frac{1}{2} m \dot{\boldsymbol{\zeta}}^{\boldsymbol{T}} \boldsymbol{S}^T \boldsymbol{S} \dot{\boldsymbol{\zeta}} + \frac{1}{2} m \boldsymbol{\zeta}^{\boldsymbol{T}} \boldsymbol{S}^T \boldsymbol{A} \boldsymbol{S} \boldsymbol{\zeta} + U_N^{(0)} = \sum_{i=1}^{3N} \frac{1}{2} [m \dot{\zeta}_i^2 + \lambda_i \zeta_i^2] + U_N^{(0)}. \tag{26.72}$$

26.6 Appendix F: Laplace transforms

Theorem: For any two functions $f(t)$ and $g(t)$ with Laplace transforms $\tilde{f}(s)$ and $\tilde{g}(s)$ we have,

$$\mathcal{L}(f * g) = \tilde{f}(s)\tilde{g}(s), \tag{26.73}$$

where $f * g$ denotes convolution integral.

Proof: Start from,

$$\mathcal{L}(f * g) = \int_0^\infty (f * g)(t) e^{-st} dt = \int_0^\infty dt e^{-st} \int_0^t f(t-u) g(u) du. \tag{26.74}$$

Change the order of integration in the r.h.s. (see figure below),

$$\int_0^\infty \int_0^t f(t-u) g(u) e^{-st} = \int_0^\infty \int_u^\infty f(t-u) g(u) e^{-st} dt du. \tag{26.75}$$

Now, change variable in the inner integral, $v = t - u$, $dv = dt$ (with u constant),

$$\int_0^\infty \int_0^\infty f(v) g(u) e^{-sv+u} dv du = \int_0^\infty f(v) e^{-sv} dv \int_0^\infty g(u) e^{-su} du = \tilde{f}(s)\tilde{g}(s). \tag{26.76}$$

The Fig. (26.1) illustrates the order of integrations in the Eq. (26.76).

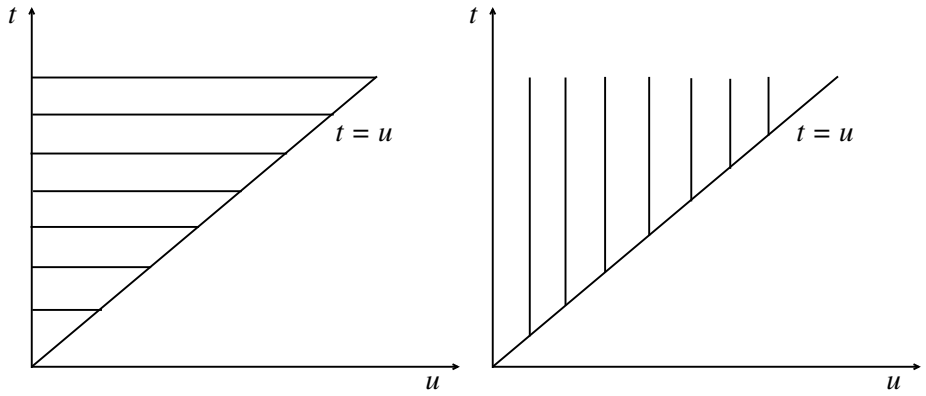


Fig. 26.1: Changing the order of integration to prove that the Laplace transform of the convolution integral is the product of Laplace transforms of the functions involved in the convolution.

26.7 Appendix G: Dirac delta function

Since the Dirac delta function and its properties are used frequently throughout this book, we find it pertinent to include a brief appendix presenting some key definitions and properties. The Dirac delta function, commonly denoted as $\delta(r)$, depends on the variable r and is defined such that it equals zero for all values of r except at $r = 0$, where it exhibits a singularity. Its defining property is that the integral over the entire real line is equal to one. Thus, the delta function can be expressed as,

$$\delta(r) = \begin{cases} \infty, & \text{if } r = 0 \\ 0, & \text{if } r \neq 0 \end{cases}, \tag{26.77}$$

which satisfies the condition,

$$\int_{-\infty}^{\infty} dr\delta(r) = 1, \tag{26.78}$$

where, if $f(r)$ is a continuous function over the interval $[a,b]$ containing the point $r = 0$, then,

$$\int_{a}^{b} drf(r)\delta(r) = f(0). \tag{26.79}$$

More generally, the symbol $\delta(r-a)$ represents the delta function centered at meaning it is zero everywhere except at $r = a$. If $r = a$ lies within the integration limits, then,

$$\int drf(r)\delta(r-a) = f(a), \tag{26.80}$$

for any continuous function $f(r)$.

Properties

The delta function satisfies the following scaling property for a non-zero scalar α,

$$\int_{-\infty}^{\infty} \delta(\alpha r)\, dr = \int_{-\infty}^{\infty} \delta(u)\, \frac{du}{|\alpha|} = \frac{1}{|\alpha|}, \tag{26.81}$$

and so

$$\delta(\alpha r) = \frac{\delta(r)}{|\alpha|}. \tag{26.82}$$

In particular, the delta function is an even distribution (symmetry), in the sense that

$$\delta(-r) = \delta(r), \tag{26.83}$$

which is homogeneous of degree -1. The distributional product of δ with r is equal to zero, $r\delta(r) = 0$. More generally, $(r-a)^n\delta(r-a) = 0$ for all positive integers n. Conversely, if $rf(r) = rg(r)$, where f and g are distributions, then,

$$f(r) = g(r) + c\delta(r), \tag{26.84}$$

for some constant c.

26.8 Appendix H: Virial expansion

The virial equation of state expresses the deviations from the ideal behavior as a infinite power series in ρ as follows,

$$\frac{P}{k_B T} = \rho\left[1 + \sum_{j=1}^{\infty} \rho^j B_{j+1}\right], \tag{26.85}$$

where $B_2(T)$, $B_3(T)$, ... are the second, third, ... virial coefficients, respectively. These virial coefficients depend solely on temperature as mentioned in the principal text. The j-th virial coefficient can be calculated in terms of the interactions of j molecules. It is convenient to derive necessary result by considering the grand canonical partition function,

$$\Xi(\mu, V, T) = \sum_{N=0}^{\infty} Q_N(N, V, T)\lambda^N, \tag{26.86}$$

where $\lambda = \exp(\mu/k_B T)$ is the activity and Q_N is the canonical partition function. Also recall that,

$$Q_N = \frac{1}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2} Z_N, \qquad Z_N = \int d\mathbf{r}_1 \cdots d\mathbf{r}_N e^{-\beta U_N(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)}, \tag{26.87}$$

cf. Eq. (12.28) for the normalized configurational partition function, and where in the absence of inter-particle interaction,

$$Q_1 = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} V; \qquad Q_N = \frac{1}{N!}\left(\frac{Q_1}{V}\right)^N Z_N. \tag{26.88}$$

Then, in accordance with Eq. (11.32),

$$PV = k_B T \ln \Xi = k_B T \ln\left(1 + \sum_{N=1}^{\infty} Q_N(V, T)\lambda^N\right). \tag{26.89}$$

Substituting Eq. (26.88) into Eq. (26.89) one obtains,

$$PV = k_B T \ln\left(1 + \sum_{N=1}^{\infty} \frac{1}{N!} z^N Z_N(V, T)\right), \tag{26.90}$$

where the notation $z = \lambda Q_1/V$ has been introduced. Expansion of $\ln(1+x)$ into the series according to Eq. (26.41) leads to the expression for the pressure,

$$P = k_B T \sum_{j=1}^{\infty} b_j z^j, \tag{26.91}$$

where the first coefficients b_j are,

$$b_1 = \frac{1}{1!V} Z_1 = 1, \tag{26.92}$$

$$b_2 = \frac{1}{2!V}(Z_2 - Z_1^2), \tag{26.93}$$

$$b_3 = \frac{1}{3!V}(Z_3 - 3Z_2 Z_1 + 2Z_1^3), \tag{26.94}$$

$\cdots\cdots$

On the other hand, the average number of particles in the grand canonical ensemble is defined by,

$$< N >= k_B T \left(\frac{\partial \ln \Xi}{\partial \mu}\right)_{V,T} = \lambda \left(\frac{\partial \ln \Xi}{\partial \lambda}\right)_{V,T}. \tag{26.95}$$

Therefore,

$$\begin{aligned} \rho = \frac{< N >}{V} &= \frac{\lambda}{V}\left(\frac{\partial \ln \Xi}{\partial \lambda}\right)_{V,T} = \frac{z}{V}\left(\frac{\partial \ln \Xi}{\partial z}\right)_{V,T} = \frac{z}{k_B T}\left(\frac{\partial P}{\partial z}\right)_{V,T}, \\ &= \sum_{j=1}^{\infty} j b_j z^j, \end{aligned} \tag{26.96}$$

where the last line follows from the application of Eq. (26.91). Now, we wish to invert this relationship, $\rho(z)$, and obtain the function $z(\rho)$. The series for the activity z is,

$$z = a_1 \rho + a_2 \rho^2 + a_3 \rho^3 + \cdots, \tag{26.97}$$

whereas from Eq. (26.96),

$$\rho = b_1 z + 2 b_2 z^2 + 3 b_3 z^3 + \cdots. \tag{26.98}$$

Then,

$$\rho = \rho(b_1 a_1) + \rho^2(b_1 a_2 + 2 b_2 a_1^2) + \rho^3(b_1 a_3 + 4 b_2 a_1 a_2 + 3 b_3 a_1^3) + \cdots. \tag{26.99}$$

Comparing the coefficients at the same degree of density on both sides of this equation, one obtains,

$$\begin{aligned} &b_1 a_1 = 1 \qquad \rightarrow \qquad a_1 = 1/b_1, \\ &b_1 a_2 = -2 b_2 a_1^2 \rightarrow a_2 = -2 b_2, \qquad b_2 = -a_2/2, \\ &b_1 a_3 = -4 b_2 a_1 a_2 - 3 b_3 a_1^3 \rightarrow a_3 = 8 b_2^2 - 3 b_3, \qquad b_3 = \frac{1}{3}(2 a_2^2 - a_3). \end{aligned} \tag{26.100}$$

Then, taking into account Eq. (26.92), the coefficients a_i are,

$$\begin{aligned} &a_1 = 1, \\ &a_2 = -2 b_2, \\ &a_3 = -3 b_3 + 8 b_2^2, \\ &\cdots\cdots \end{aligned} \tag{26.101}$$

Thus the function $z(\rho)$ can be written as,

$$z = \rho - 2 b_2 \rho^2 + (8 b_2^2 - 3 b_3)\rho^3 + \cdots. \tag{26.102}$$

We can substitute this result into Eq. (26.91) to get,

$$\begin{aligned}\frac{P}{k_BT} =& b_1z + b_2z^2 + b_3z^3 + \cdots = z - \frac{1}{2}a_2z^2 + \frac{1}{3}(2a_2^2 - a_3z_3) + \cdots \\ =& (a_1\rho + a_2\rho^2 + a_3\rho^3 + \cdots) - \frac{a_2}{2}(a_1\rho + a_2\rho^2 + a_3\rho^3 + \cdots)^2 + \\ & \frac{1}{3}(2a_2^2 - a_3)(a_1\rho + a_2\rho^2 + a_3\rho^3 + \cdots), \\ =& a_1\rho + \rho^2\left(a_2 - \frac{a_1a_2}{2}\right) + \rho^3\left(a_3 - a_1a_2^2 + \frac{1}{3}(2a_2^2 - a_3)a_1^3\right) + \cdots \\ =& \rho + \rho^2(-b_2) + \rho^3\left(\frac{1}{3}(2(8b_2^2 - 3b_3) - 4b_2)\right) + \cdots \\ =& \rho + \rho^2(-b_2) + \rho^3(4b_2^2 - 2b_3) + \cdots,\end{aligned} \tag{26.103}$$

or equivalently,

$$\frac{P}{k_BT} = \rho + B_2(T)\rho^2 + B_3(T)\rho^3 + \cdots, \tag{26.104}$$

where

$$B_2(T) = -b_2 = -\frac{1}{2V}(Z_2 - Z_1^2), \tag{26.105}$$

$$B_3(T) = 4b_2^2 - 2b_3 = \frac{1}{V^2}(Z_2 - Z_1^2) - \frac{1}{3V}(Z_3 - 3Z_2Z_1 + 2Z_1^3), \tag{26.106}$$

where $B_2(T)$ and $B_3(T)$ are the second and third virial coefficients, respectively.

We must now show that these results are equivalent to the Mayer expansion results obtained in the canonical ensemble. For this, we first calculate the second and third virial coefficients, we need,

$$Z_1 = \int_V d\mathbf{r}_1 = V, \tag{26.107}$$

$$Z_2 = \int_V d\mathbf{r}_1 d\mathbf{r}_2 e^{-\beta u(r_{12})} = \int_V d\mathbf{r}_1 d\mathbf{r}_2 (1 + f_{12}) = V^2 + \int_V d\mathbf{r}_1 d\mathbf{r}_2 f_{12}, \tag{26.108}$$

and

$$Z_3 = \int_V d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 e^{-\beta[u(r_{12}) + u(r_{13}) + u(r_{23})]}. \tag{26.109}$$

where $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$. Then, $B_2(T)$ in Eq. (26.105) is,

$$\begin{aligned}B_2(T) &= -b_2 = -\frac{1}{2V}(Z_2 - Z_1^2), \\ &= -\frac{1}{2V}\int_V d\mathbf{r}_1 d\mathbf{r}_2 f_{12}(r_{12}),\end{aligned} \tag{26.110}$$

In order to evaluate B_3, we need,

$$\begin{aligned}Z_3 &= \int_V d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 (1 + f_{12})(1 + f_{13})(1 + f_{23}), \\ &= \int_V d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \left[f_{12}f_{13}f_{23} + f_{12}f_{13} + f_{12}f_{23} + f_{13}f_{23} + f_{12} + f_{13} + f_{23} + 1\right],\end{aligned} \tag{26.111}$$

as well is Z_2 and Z_1. Then, B_3 from Eq. (26.106) is,

$$B_3(T) = -\frac{1}{3V}\int\int\int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 f_{12}f_{13}f_{23}. \tag{26.112}$$

Bibliography

[1] L. García-Colín, *Termodinámica estadística*. Universidad Autónoma Metropolitana - Unidad Iztapalapa, 1995.

[2] H. T. Davis, *Statistical Mechanics of phases, interfaces, and thin films*. Wiley-VCH, Inc., 1996.

[3] A. Davydov, *Quantum Mechanics*. International series in natural philosophy, Pergamon Press, 1976.

[4] I. Vakarchuk, *Quantum Mechanics*. Lviv State University (In Ukrainian), 1998.

[5] V.B. Kobyliansky, *Statistical Physics.* Kiev, Vyshcha Shkola, 1972.

[6] I. Yukhnovsky and M. Holovko, *Statistical Theory of Classical Equilibrium Systems*. Kiev, Naukova Dumka (In Russian), 1980.

[7] D. McQuarrie, *Statistical Mechanics*. Harper and Row, Publishers, Inc., 1973.

[8] D. McQuarrie, *Mathematical Methods for scientists and Engineers*. University Scence Books, 2003.

[9] F. del Río, *Conceptos de la física: I modelos clásicos*. Alhambra Mexicana, 1986.

[10] L. D. Landau and E. M. Lifshitz, *Statistical Physics*. Reading, MA: Addison-Wesley, 1958.

[11] E. de Miguel and C. Vega, "The global phase diagram of the Gay-Berne model," *J. Chem. Phys.*, vol. 117, pp. 6313–6322, 2002.

[12] G. N. Clark, A. J. Haslam, A. Galindo, and G. Jackson, "Developing optimal wertheim-like models of water for use in statistical associating fluid theory (SAFT) and related approaches," *Mol. Phys.*, vol. 104, pp. 3561–3581, 2006.

[13] V. M. Trejos, O. Pizio, and S. Sokolowski, "On the theoretical description of the liquid-vapor coexistence of water-like models with square-well attraction and site-site chemical association," *Fluid Phase Equil.*, vol. 473, pp. 145–153, 2018.

[14] W. A. Steele, "The physical interaction of gases with crystalline solids. I. Gas-solid energies and properties of isolated adsorbed atoms," *Surf. Sci.*, vol. 36, pp. 317–352, 1973.

[15] W. A. Steele, *The Interaction of Gases with Solid Surfaces*. Oxford: Pergamon, 1974.

[16] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*. Oxford University Press, 1989.

[17] D. Nicholson and G. Parsonage, *Computer Simulation and the Statistical Mechanics of Adsorption*. Academic Press, 1982.

[18] R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. 2nd edn. John Wiley & Sons, Inc., New York., 1985.

[19] R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics*. John Wiley & Sons, Inc., New York., 1975.

[20] M. Preston and R. Bhaduri, *Structure of the Nucleus*. Addison-Wesley, London., 1975.

[21] D. J. Griffiths, *Introduction to Quantum Mechanics*. Pearson Prentice-Hall, second edition, New Jersey., 2005.

[22] C. E. Moore, *Atomic energy levels*. US Department of Commerce, Natl. Bur. Standards, Vol. 1., 1949.

[23] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Saunders College, New York, 1976.

[24] W. E. Henry, "Spin paramagnetism of Cr^{+++}, Fe^{+++}, and Gd^{+++} at liquid helium temperatures and in strong magnetic fields," *Phys. Rev.*, vol. 88, pp. 559–562, 1952.

[25] F. Reif, *Fundamentals of statistical and thermal physics*. Mc Graw Hill, New York, 1965.

[26] H. E. Stanley, *Introduction to phase transitions and critical phenomena*. Oxford University Press: New York., 1987.

[27] P. Atkins and J. D. Paula, *Atkins' Physical Chemistry*. W. H. Freeman and Company, New York, Eighth Edition, 2006.

[28] T. van Erp and J. Martens, "A standardization for BET fitting of adsorption isotherms," *Microporous and Mesoporous Mater*, vol. 145, pp. 188 – 193, 2011.

[29] M. Jaroniec and M. Kruk, "Standard nitrogen adsorption data for characterization of nanoporous silicas," *Langmuir*, vol. 15, pp. 5410–5413, 1999.

[30] R. Durrer, *The Cosmic Microwave Background*. Cambridge University Press., 2008.

[31] H. T. Davis, "Density distribution functions of confined Tonks-Takahashi fluids," *J. Chem. Phys.*, vol. 93, pp. 4339–4344, 1990.

[32] J. A. Cuesta and A. Sánchez, "General non-existence theorem for phase transitions in one-dimensional systems with short range interactions, and physical examples of such transitions," *J. Stat. Phys.*, vol. 115, pp. 869–893, 2004.

[33] P. Monson, "The properties of inhomogeneous square-well mixtures in one dimension," *Mol. Phys.*, vol. 70, pp. 401–423, 1990.

[34] G. Rutkai, M. Thol, R. Span, and J. Vrabec, "How well does the Lennard-Jones potential represent the thermodynamic properties of noble gases?," *Mol. Phys.*, vol. 115, pp. 1104–1121, 2017.

[35] T. Hill, *Statistical Mechanics: Principles and Selected Applications*. Dover Books on Physics, McGraw-Hill, 1956.

[36] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids*. John Wiley & Sons, Inc., 1964.

[37] M. S. Wertheim, "Fluids with highly directional attractive forces. I. Statistical thermodynamics," *J. Stat. Phys.*, vol. 35, no. 1, pp. 19–34, 1984.

[38] M. S. Wertheim, "Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations," *J. Stat. Phys.*, vol. 35, no. 1, pp. 35–47, 1984.

[39] M. S. Wertheim, "Fluids with highly directional attractive forces. III. Multiple attraction sites," *J. Stat. Phys.*, vol. 42, no. 3, pp. 459–476, 1986.

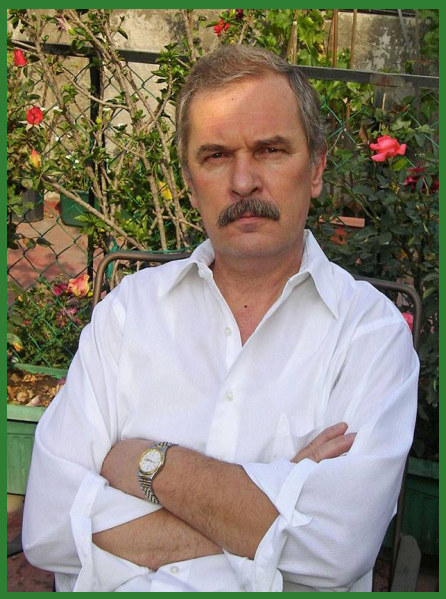
[40] M. S. Wertheim, "Fluids with highly directional attractive forces. IV. Equilibrium polymerization," *J. Stat. Phys.*, vol. 42, no. 3, pp. 477–492, 1986.

[41] P. T. Cummings and G. Stell, "Statistical mechanical models of chemical reactions. analytic solution of models of, $a + b \rightleftharpoons ab$ in the Percus-Yevick approximation," *Mol. Phys.*, vol. 51, pp. 253 – 287, 1984.

[42] H. L. Lemberg and F. H. Stillinger, "Central-force model for liquid water," *J. Chem. Phys.*, vol. 62, pp. 1677–1690, 1975.

[43] P. T. Cummings and G. Stell, "Statistical mechanical models of chemical reactions II. Analytic solution of models of the Percus-Yevick approximation for a model of homogeneous association," *Mol. Phys.*, vol. 55, pp. 33–48, 1985.

[44] P. T. Cummings and G. Stell, "Statistical mechanical models of chemical reactions: III. Solvent effects.," *Mol. Phys.*, vol. 60, pp. 1315–1342, 1987.



Dr. Orest Pizio was born in 1952 at the city of Lviv in Ukraine. He graduated with honors from the Department of Physics of the University of Lviv in 1973. He worked at Bogolyubov Institute of Theoretical Physics (ITP) of the National Academy of Sciences of Ukraine (NASU) and Yukhnovskii Institute for Condensed Matter Physics (ICMP) of NASU during 1973 – 1993 starting from temporary young researcher up to senior research fellow. He received his Ph.D. in 1980 at the ITP of NASU under supervision of Prof. I.R. Yukhnovskii and Prof. M.F. Holovko. Since 1993 Dr. Orest Pizio is working at the Institute of Chemistry of UNAM. He is a national researcher emeritus of SNII of Mexico, and Dr. Honoris Causa of the ICMP of NASU. Orest Pizio has authored more than 250 research articles and several chapters in collective monographs. He served to scientific community as a reviewer, editor of special issues and as an organizer of international congresses. His studies are focused in the statistical theory and in computer simulations of liquids, liquid mixtures and electrolyte solutions, as well as in non-homogeneous associating fluids.

Dr. Víctor Manuel Trejos was born in Riosucio, Caldas, Colombia. He received his B.Sc. in Chemical Engineering and his M.Sc. degree from the National University of Colombia. In 2014, he obtained his Ph.D. in Physics from the Institute of Physics at the University of Guanajuato, Mexico, under the supervision of Prof. A. Gil-Villegas. Following his doctoral studies, he held postdoctoral research positions at Vanderbilt University under the supervision of Prof. Peter T. Cummings and at the National Autonomous University of Mexico (UNAM) under the supervision of Prof. Orest Pizio. Since 2023, Dr. Trejos has been an Assistant Professor in the Department of Chemistry at the Autonomous Metropolitan University-Iztapalapa (UAM-Izt) in Mexico City, Mexico. Dr. Trejos has authored more than 60 scientific publications and has supervised over 15 undergraduate and graduate students. His research interests focus on the study of inhomogeneous fluids and aqueous ionic systems through density functional theory (DFT), perturbation theory, and molecular simulations.