# Equilibrium Statistical Mechanics (Hypertext version May 7, 2012)[1]

This version has the last chapter that gives all the solutions to the chapter-end problems.

Yoshi Oono yoono@uiuc.edu
Department of Physics and Institute of Genomic Biology
University of Illinois at Urbana-Champaign

---

[1] Problems and their solutions are not yet revised from the February version.

To the memory of Bob Clegg (July 18, 1945 - October 15, 2012),

sharing critical points of view with him was extremely encouraging.....

This is a serious introductory textbook of equilibrium statistical mechanics. Its main target readers are:
1) those who have already learned undergraduate level thermal physics,[2] but have not been satisfied,
2) those who wish to review elementary statistical mechanics without returning to introductory textbooks,
3) those who do not wish to be a mere consumer of statistical mechanics/those who have ambition to develop their own viewpoints (or even to make their own theoretical frameworks).
The author wishes to make a textbook of equilibrium statistical mechanics for those who wish to think things critically by themselves.

For the author to understand a theory is to reconstruct it by himself, or to make a textbook of the theory for himself since his student days (thus most lectures were very hard to understand during the class). The author tries to make a textbook that he wished to read when he was a student. Those who review statistical mechanics for exams' sake are not among the target readers of this book; those people should think that the life after exams is much longer. For the life after exams the reader need not be able to solve exercise problems quickly; in real life if the reader wishes to solve them, s/he can simply look up the problem books. Being tenacious is much more important than being quick and smart in order to be a real scientist.

**However, the author has no intention to make a book quite different from the standard introductory textbooks, so all the elementary topics are covered in detail (almost from the ground level).** Furthermore, to make the book reasonably inexpensive, there is a supplementary web page called "Equilibrium Statistical Mechanics Supplementary Pages," which covers detailed or subtle discussions and advanced topics (with detailed calculation). The author was an organic synthetic chemist up to the master's degree, and has not attended any standard physics (and mathematics) courses. He learned statistical mechanics by himself, solving Kubo's exercise book[3] that contains full solutions of all the problems. Therefore, to help those who struggle by themselves all the detailed solutions to the problems at the end of each chapter are also posted at the Supplementary Pages. Consequently, the book should be useful for those who wish to review equilibrium statistical mechanics for the sake of exams as well.

---

[2] At UIUC the 200 level courses.

[3] R. Kubo, H. Ichimura, T. Usui, and N. Hashitsume, *Statistical Mechanics* (North-Holland, 1990).

One could say that statistical mechanics is a wonderful tool, so we ordinary scientists should become expert users without ever asking why it works. It is true that axioms cannot be derived. However, even in mathematics it is not useless nor futile to understand why particular axioms are adopted. Furthermore, it is against the spirit of science to swallow what is told without critical thinking. Therefore, the book does not subscribe to the opinion mentioned above; the author wishes to make a textbook that might be useful to those who wish to make 'better' statistical mechanical framework, or statistical framework applicable to the systems beyond the scope of the conventional theories. Therefore, the book discusses not only how to use statistical mechanics but also why its theory is plausible and natural. The book places certain stress on the general idea of statistical mechanics and its consequences.

In the ordinary introduction to our subject, it is not rare that the students are persuaded to become able to solve problems without trying to understand why s/he can solve them. This is not at all a bad attitude, *if* they will later start to think, so this book discusses clever techniques as well. Still, the author wishes to treat physics as an intellectual and cultural endeavor.[4] Therefore, most explanations will be with informal backgrounds that exhibit intuitive ideas [Spilt over explanations will be in the Supplementary Pages].

There are folklores in statistical mechanics. For example, in many textbooks ergodic theory and the mechanical foundation of statistical mechanics are discussed even though detailed mathematical explanations may be missing. We must clearly recognize such topics are almost irrelevant to statistical mechanics. We are also brainwashed that statistical mechanics furnishes the foundation of thermodynamics, but we must clearly recognize that without thermodynamics statistical mechanics cannot be formulated. It is a naive idea that microscopic theories are always more fundamental than macroscopic phenomenology. Since natural science must be empirically supported, phenomenology is the most reliable part of science (its least metaphysical portion[5]). Many physicists, including the author himself, who wish to emphasize theory and mathematical structures, tend to believe the so-called fundamental (microscopic) theories are the foundation of foundations, but occasionally we

---

[4] Since mathematics is an integral part of our intellectual culture, it will be used without any reserve. However, in most cases, some advanced topics will be with footnotes and supplementary notes in the web or small lettered auxiliary explanations that should be accessible without difficulty by physics students.

[5] Crudely put, what we cannot decide experimentally is metaphysics.

should reflect on the premises.

Topics that we will cover but that may not be discussed in detail in many textbooks include the following topics:
∗ Intuitive explanations are not avoided, or scaffolds will not be removed from theories. For example, a physical interpretation of the so-called ensemble theory is provided.
∗ The pillars of the modern probability theory are the law of large numbers and its refinements (large deviation theory and central limit theorem). The pillars of statistical mechanics are closely related to them. The law of large numbers makes statistical mechanics possible. Fluctuation theory is the study of large deviation, and renormalization group theory is the central limit theorem.
∗ Legendre transformation (actually the Fenchel equalities are called Legendre transformations in elementary textbooks) is important transformations in thermodynamics. It is a fundamental concept in convex analysis. The equivalence of ensembles is closely related to the self-dual nature of Legendre transformations.
∗ Today, there must be a lot of opportunities to learn quantum many-body theory, but it may be rare that physics students are exposed to classical liquid. Thus, imperfect gas, simple liquid, and some needed tools to study them are outlined.
∗ In conjunction to phase transitions, introductory discussions on thermodynamic limit and Gibbs states are given. More generally, the book tries to make mathematical results more easily accessible than the original papers and mathematical physics textbooks. Important mathematical results often used in statistical mechanics (e.g., Stirling's formula, Jensen's inequality, Perron-Frobenius' theorem) are stated with proofs or theoretical-physicist-style formal arguments.
∗ The arrangement of the topics is probably more logical than many textbooks. For example, most introductory textbooks discuss mean field theory first, and then state that unfortunately ubiquitous thermal fluctuation makes it not very reliable. This book stresses the importance of equilibrium fluctuations. Inevitably, the idea of renormalization comes to the fore. Only after renormalization can one understand the significance of mean field theory in statistical mechanics. Thus, renormalization is outlined before mean field approaches.

The book consists of the main text and (indented) additional explanations in small letters and footnotes. To understand the main text no particular reading of the other parts is required. A reasonable outline of statistical mechanics should be understood by reading only the main text. Small lettered explanations cover reviews of related elementary topics and advanced topics (mostly with bold face titles). The explanations with † are more advanced. Long footnotes are often with titles embraced by ⟪ ⟫. These footnotes augment related topics, some of which may be read as short column articles. As stated above, the book has Supplementary Pages that contain more advanced topics, explanation of subtle points, and summary of each section.

Each chapter ends with a set of problems most of which have already been asked as homework and exam problems. Most representative types of problems in problem books may be found there. Even if the reader cannot do them quickly, that is perfectly all right, provided s/he continues to think tenaciously. As already mentioned, all the full solutions are found in Supplementary Pages.

The book is based on what the author learned from his old friends, especially, Bob Clegg, Hisao Hayakawa, Shin-ichi Sasa, Ken Sekimoto, Akira Shimizu, Yoichiro Takahashi and Hal Tasaki, although he has not been a very good student. Extensive comments of Joel Cannon, Barry Friedman, and Chuck Yeung have been incorporated. The precursor of this book is a posted set of lecture notes for the graduate statistical mechanics course of University of Illinois at Urbana-Champaign. Corrections and suggestions to the notes due to Hyeong Jun Kim, Farzad Tehranchi, Youcai Wang and Shizhong Zhang are gratefully incorporated.

On September 20th, 2006, Akira Shimizu organized a 'town meeting' on the manuscript in Tokyo. There, the author obtained critical comments of not only Shimizu and Tasaki already mentioned above but also Takashi Hara and Nobuyasu Itoh. This version has supposedly taken care of all the complaints and criticisms raised in the meeting. The author is grateful to the participants to the town meeting. Further later, he had an opportunity to discuss with Hara, and realized several wrong statements in Chapter 5. The second town meeting organized by Hisao Hayakawa was held on December 20, 2006 in Kyoto. In this meeting the author got many requests of incorporating various topics, but regretfully the version has not responded to most of them because the thickness of the book cannot be enormous. The 2008 March Japanese version with more introductory materials was circulated and the author got again extensive criticisms by Tasaki, Shimizu, and Takahashi. They have been duly incorporated into this version. On Jan 7-9. 2009 an opportunity of the concentrated course was provided by Kiyoshi Higashijima. Useful comments and questions by Toshio Nakatsu and Makoto Kikuchi at this opportunity and by Akira Yoshimori at a similar opportunity at Kyushu University in June 2009 provided by Hiizu Nakanishi have been gratefully incorporated. The resultant 2009 Japanese version was further subjected to Sasa, Shimizu and Tasaki's criticisms. The results of discussions at the meetings with them at University of Tokyo in early July, 2010 have been take care of.

**Special symbol for hyper-referencing: 'jump-mirror system'**

Some concepts in this color are buttons to jump to related topics/explanations, where there is this mark **r**. Clicking this mark, you can return to your starting point. You can test. This version with full solutions as Chapter 6 has an analogous mark **q** that returns the reader to the problems or their solutions.

## The Supplementary Web Page

`http://www.yoono.org/Y_OONO_official_site/StatPhys`

In order not to make the book too thick, a summary of each section (minimum), detailed explanation and auxiliary comments, examples, more advanced topics and detailed solutions to problems at the end of the chapters are posted here. For example, requested materials such as the $\varepsilon$-expansion, a method to obtain the exact free energy of the square lattice Ising model are posted here.

**r** allows you to jump back to 'test.'

# Contents

# Chapter 1

# Basics

For a practical physicist equilibrium statistical mechanics implies the following general ‘algorithm’: Take a many-body system whose Hamiltonian is $H$.[1] Then, the Helmholtz free energy $A$ of the system at absolute temperature $T$ is given by (for a brief review see “work and free energy” below)

$$A = -k_BT\log Z, \tag{1.0.1}$$

where[2] $k_B$ ($= 1.3806503\times 10^{-23}$ J/K) is the Boltzmann constant and $Z$ is the canonical partition function computed as

$$Z = Tr\, e^{-\beta H}. \tag{1.0.2}$$

Here, $\beta = 1/k_BT$ (a standard notation) and the trace implies quantum-mechanically the sum over all the eigenstates of the Hamiltonian, and classical-mechanically the integration over the phase space. Thus, the algorithm supplied by statistical mechanics allows us to compute thermodynamic quantities of a system from its Hamiltonian that is a microscopic description of the system. This is the theoretical core of statis-

---

[1] To enjoy this book, the reader should know elementary analytical mechanics (the canonical equation of motion, Poisson brackets, etc.), elementary quantum mechanics (bra and ket, angular momentum, etc.) and elementary thermodynamics. For these elementary subjects appropriate summaries will be given in Supplementary Pages. However, the reader may have forgotten thermodynamics, so rudiments of equilibrium thermodynamics will be added at appropriate places.

[2] In this book ‘ln’ will never be used; log always implies the natural logarithm.

tical mechanics, and our remaining task is only to evaluate $Z$....(canonical formalism)[3]

The purpose of the book is to help the reader understanding the above theoretical core of statistical mechanics, and then becoming able to use its machinery.[4]

The algorithm summarized at the beginning of the Chapter is often called the *ensemble theory* as some of the readers may already know, but what ensemble does it concern with? Perhaps, the reader has heard furthermore about ergodic hypothesis that equates the time average and the average over the allowed set of microscopic states. Is this hypothesis really relevant to statistical mechanics?

The main discussion of statistical mechanics will begin at Sect. 1.3. Section 1.1 may be read as the "making of statistical mechanics," and Sect. 1.2 is a review of probability theory and the law of large numbers. Section 1.1 is addressed to those who have already learned elementary statistical mechanics. Therefore,

**if the reader wishes to be efficient, s/he can start from Sect. 1.3**.

Or, the reader can start from Sect. 1.3, read up to Sect. 1.6, and then come back to Sect. 1.1. The author is acutely aware that there are people who denounce attempts to discuss 'making of statistical mechanics.' The relation between thermodynamics and equilibrium statistical mechanics is a paragon of the relation between a phenomenology and its statistical model. Therefore, the author believes that to understand intuitively why the relation becomes plausible should be of some use, when one wishes, e.g., to extend the statistical framework to systems away from equilibrium or to the systems the traditional framework has not been applied (or when one wishes to understand why such extensions are impossible). This is the reason why we dare try to see the "making of statistical mechanics."

This book has fine-lettered entries with titles such as the following. They may be reviews of some elementary topics or slightly advanced or complicate materials. Roughly speaking, if the reader knows elementary thermodynamics and mechanics, without reading these fine-lettered entries, s/he can acquire general understanding of statistical mechanics and a certain extent of technical details.

**Work and free energy**
The readers are expected to be familiar with the outline of thermodynamics, but many of them may have forgotten it (as already noted in footnote 4, an introductory material is provided in Supplementary Pages). Therefore, in this Chapter rudiments

---

[3] However, it is usually hard to compute $Z$, so statistical mechanics consists of three parts: the study of results obtainable without any concrete computation, the study of approximate methods to evaluate $Z$, and the study of systems for which $Z$ can be computed exactly.

[4] **《Introduction to Statistical Thermodynamics》** An introductory material for those who have never learned thermodynamics and statistical mechanics at all is posted as a part of Supplementary Pages. This is essentially a set of undergraduate course lecture notes. These notes provide more than enough elementary material for this book.

of equilibrium thermodynamics are summarized at strategic points in small letters (like this paragraph) marked with [T]. An outline of the first and the second laws is in Sect. 1.3, but here a common-sense understanding of free energy is summarized briefly. The precise meaning of 'work' and 'heat' will be discussed later, but here they may be interpreted intuitively.

In thermodynamics, the energy conservation law is summarized as the first law of thermodynamics: the change of internal energy $\Delta E$ consists of heat $Q$ and work $W$:

$$\Delta E = Q + W. \tag{1.0.3}$$

The sign convention is: if a quantity goes into (resp., out from) the system, it is $+$ (resp., $-$). For example, if we do work to the system, $W > 0$; if the system cools by radiating heat, $Q < 0$; etc. For a reversible isothermal change, we may write $Q = T\Delta S$, where $\Delta S$ is the change of the system entropy, and $T$ is the absolute temperature of the system. For a reversible isothermal[5] process we may write

$$W = \Delta(E - TS) = \Delta A, \tag{1.0.4}$$

introducing the Helmholtz free energy $A = E - TS$. This is the meaning of the statement that the reversible work added to the system through an isothermal process is equal to the free energy increase of the system.

One way to state the second law of thermodynamics is that the free energy change of the system does not exceed the work added to the system (Thomson's principle) for an isothermal process:

$$\Delta A \le W. \tag{1.0.5}$$

This inequality holds even when $W$ is the work done by the system (say, to us). In this case, $W < 0$ (consequently, $\Delta A < 0$ as well). That is, in magnitudes, $|\Delta A| \ge |W|$: the work we can take out from the system cannot exceed the free energy decrease of the system. □

## 1.1 Being macroscopic—why do we need statistics?

Statistical mechanics *assumes* that a macroscopic object (a macrosystem) is made of numerous 'microscopic constituents' (particles) obeying (a certain) mechanics. Informally, a 'macroscopic object' implies an object of our length scale (say, not far smaller

---

[5] ⟪**Isothermal process**⟫ An isothermal process is a process that can take place in a system thermally contacting with a constant temperature heat bath. The process need not be reversible, so during the process temperature may not be definable, but its initial and final equilibrium states must have the same temperature. Needless to say, for a reversible isothermal process the system temperature is always kept constant.

than 0.1mm but not far larger than 10 m, i.e., $10^{\pm 3}$cm) that is sufficiently uniform. 'Microscopic constituents' are, usually, molecules and atoms. Mechanics means the fundamental laws of mechanics (classical or quantum mechanics at present). The microscopic state (*microstate*) of a macrosystem is its individual state that can be completely specified by its mechanical description (i.e., the elementary states of a system described as a mechanical system); for a classical system it is a point in the phase space of the system, and for a quantum system it is a 1-dimensional subspace (called a *ray*) of the state (complex) vector space (Hilbert space) of the system. There are so many microscopic constituents that if we wish to explain macroscopic observables in terms of the properties and behavior of microscopic constituents obeying a certain mechanics, we should need a certain statistical means. This is the usual explanation of why we need statistics, motivating 'statistical mechanics.'[6]

Is the meaning of the word 'macroscopic' clear? There might be readers who regard it not so objective to characterize the concept 'macroscopic' in conjunction to our own size. Furthermore, however large an object may be, if it is not homogeneous up to a small scale, we could not simply say it is a macroscopic object. After all, it is not an easy task to specify the word 'macroscopic' generally. Therefore, here, we consider this only within the context of equilibrium statistical mechanics.

To begin with what is 'equilibrium'? Let us recall the fact summarized as the so-called *zeroth law of thermodynamics* (see [T1]). If a 'macroscopic' object left for a sufficiently long time[7] without any interaction with its surrounding world (i.e., left alone in isolation[8]), it reaches a macroscopically time-independent state, called

---

[6] However, the so-called 'ensemble theory' does not take statistics over microscopic constituents, but over microstates of a macrosystem.

[7] ⟪**How long is long enough?**⟫ If no further (macroscopic) change can be observed by patient observation, the observation is long enough. Therefore, according to Feynman, an equilibrium state is, in practice, a state in which 'all the fast events have already occurred, but no slow events have started yet.' Feynman's lecture notes: *Statistical Mechanics, a set of lectures* [notes taken by R. Kikuchi and H. A. Feiveson, edited by J. Shaham] (W. A. Benjamin, Inc., 1972) is worth owning, although its level is rather advanced.

[8] ⟪**"To be left alone in isolation"**⟫**r** In plain terms, this means that the system is left without any exchange of energy, matter and information with its surrounding environment. Then, can we isolate a constant volume gas? We need a container, and the container is a part of the external environment of the gas. In thermodynamics, such as containers are idealized as walls, and are excluded from 'physics,' so a system isolated by a wall is regarded isolated, although it interacts with the wall. In statistical mechanics we must describe such systems as mechanical systems. We wish to have corresponding idealization of 'walls.' In statistical mechanics we say a system is isolated, if we can model the system in terms of mechanics with a time-independent uniform boundary condition. Notice that the key point is that an isolated system can be described in terms of mechanics.

an *equilibrium state*. However, isn't this characterization useless to our purpose? We wish to characterize the word 'macroscopic' or 'macroscopically,' but such words already appear in the above sentence. If we wish to avoid circularity, we can only accept the following answer: we say, in equilibrium, a system is macroscopic, if equilibrium thermodynamics holds sufficiently accurately. In Sect. 1.3 the meaning of this statement will be made clearer, but here the readers have only to have its rough understanding.

The zeroth law of thermodynamics [T1] further requires the following as well. If a macroscopic object in equilibrium is cut into halves, (1) each half is individually in equilibrium in isolation, and (2) if the halves are joined again, the joined result is (thermodynamically) indistinguishable from the original state of the original object (Fig. 1.1.1).

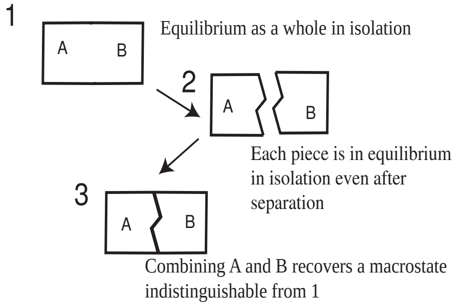


Fig. 1.1.1 Illustration of a part of the zeroth law of thermodynamics we need. The usual zeroth law asserts the existence of equilibrium states and the equilibrium relation as an equivalence relation: if systems $A$ and $B$ are in equilibrium when joined, and if $B$ and $C$ are also in equilibrium when joined, so are $A$ and $C$. However, the invariance of equilibrium states under partition and rejoining is much more fundamental.

**[T1] The zeroth law of thermodynamics**[9] (**r**to p12, **r**to [T6])
This law guarantees the existence of the equilibrium state and characterizes its nature.
(1) An isolated system (see footnote 8),[10] left alone for a sufficiently long time, reaches

---

[9] Fine-lettered explanations with [T] are about thermodynamics.

[10] ⟪**E. Borel and chaos**⟫**r** E. Borel (the creator of measure theory) pointed out long ago that the alteration of the gravitational field around us due to a 1cm displacement of 1g mass on Sirius was sufficient to change the very near future dynamics of a gas particle in front of us. There is no way to shield gravitational field, so there cannot be any truly isolated system in the universe. We know that for many mechanical systems the external noise zero limit and the long time limit do not commute. Consequently, we should recognize that the traditional idealization of statistical mechanics considering an isolated Hamiltonian system is not a good idealization of reality. The equilibrium state in thermodynamics should be understood as a state that does not change macroscopically under such unavoidable external noises (we could say under thermodynamic adiabatic conditions). This is probably the reason why it is fundamentally difficult to understand thermodynamics in terms of mechanics. Maxwell and Boltzmann thought that the influences of the external world were unavoidable; it is worth keeping in mind that they seem to have thought that equal probability sampling of microstates was possible because of external influences.

a (macroscopically) time-independent state.**r**[11] This state is called an *equilibrium state*.

(2) If the thus obtained equilibrium system is partitioned into two (approximately equal) parts (by a plane), then

(i) each piece in isolation is in equilibrium, and

(ii) if these pieces are joined as before the partition, the joined result is in equilibrium as a whole, and its state cannot be (thermodynamically) distinguished from the state before the partition.

The usual zeroth law states, instead of (2), that the thermal equilibrium relation is an equivalence relation as mentioned in the caption of Fig. 1.1.1. However, we cannot state such a requirement at the beginning without introducing the concept of thermal contact.

**Comment about 'isolated systems'** The just stated zeroth law is about the 'isolated' system not under the influence of external fields as in the standard thermodynamic textbooks. However, there are many thermodynamic states that cannot be realized when the system is isolated without any external fields. A state of a magnet at high temperatures with a nonzero magnetization is such an example. In such cases the presence of an external field necessitates exchange of energy between the system and its surrounding world. We must relax the 'isolation' condition, because we need specify conjugate variables (e.g., magnetic field) to fix (the average values of ) thermodynamic coordinates (e.g., magnetization).

A required modification will be given at the end of Section 1.3. For the time being we consider the conventional isolated systems. The required modification is, in short, to replace the isolation condition with the 'isolation condition *sensu lato*' (the condition that allows time-independent mechanical operation to the isolated system in the usual sense); henceforth the condition will be called '*thermal isolation*.' If we replace the concept 'isolation' with 'thermal isolation,' all the statements will still be correct.

**(Ferro)magnets**

To illustrate concepts and situations magnets are often used in this book, so let us review magnets and their salient phenomena. (Ferro)magnets are microscopically understood as systems in which atomic or molecular magnets (atoms and molecules with unpaired electrons) are spatially arranged and are interacting with each other. Between a pair of these elementary magnets is an interaction that tends to align their magnetic moments (the so-called exchange interaction). If the temperature is sufficiently high, elementary magnets point random directions, so there is no magnetization (= the algebraic sum of magnetic moments of elementary magnets). If the temperature is lowered, then below some temperature, which is called the *critical temperature*, the effect of interaction among elementary magnets overcomes the disordering effect of thermal noise and the elementary magnets align predominantly in a certain direction.

---

[11] This is an assertion as to the direction of time. Notice that it is NOT the second law. What is the meaning of the clause 'macroscopically time independent'? As already alluded, a macrostate is a state specified by thermodynamic observables, and 'time independence' implies that thermodynamic observables are time-independent. However, thermodynamic observables cannot be characterized without the first and the second laws (see [T5]). That is, precisely speaking, the zeroth law cannot be understood separately from other thermodynamic laws.

Consequently, the system exhibits a non-zero magnetization. The system Hamiltonian does not have any special direction to which the magnetic moments must order, but below the critical point this symmetry is lost because of self-interactions. This is called the spontaneous symmetry breaking. In some magnets their crystal lattices have special directions (the easy directions) into which elementary magnetic moments predominantly point. If there are only two opposite directions (say, + and − directions) that elementary magnetic moments can point, the magnet is called an Ising magnet. In such a magnet, below its critical temperature, the + or the − direction is pointed by majority of elementary magnets and the magnetization becomes non zero. However, whether the + phase and the − phase can coexists (just as ice and liquid water at 0 °C under 1 atm) depends on the spatial dimensionality (see Section 5.4).

If a macroscopic system in equilibrium is halved, the resultant halves are, if isolated, individually in equilibrium again. For this statement to be literally true, a macroscopic system must be infinitely large. Strictly speaking, the laws of thermodynamics are laws valid for infinite systems ('systems in the thermodynamic limit'). Precisely speaking, all the proper assertions of statistical mechanics are those after the thermodynamic limit is taken. This limit is not very simple (see Chapter 5, esp., Sections 5.2 and 5.3). However, for equilibrium thermodynamics to hold with sufficient accuracy the system need not be infinitely large; if it is about our size, the scale is more than enough. That is why thermodynamics is practically meaningful. Thermodynamic observables may be determined with considerable accuracy for systems of our size and for systems far smaller than our size, say, for 0.1 mm cubes. However, the accuracy of thermodynamic observable values is restricted by 'fluctuations' in the observables; their relative size is proportional to the inverse square root of the system size (measured in, e.g., mass). See Sect. 2.8 and the fine-lettered explanation below entitled 'The relation between thermodynamic limit and finite systems.'

Statistical mechanics bridges mechanics and thermodynamics, so let us make it clear between what sort of observables it relates. Let us assume that we have a common sense notion of mechanical observables. What is the characteristic of thermodynamic observables? Important empirical facts answering this question is summarized as the *fourth law of thermodynamics*:

> Thermodynamic observables are either intensive or extensive.[12]

**r**The fourth and the zeroth laws entail that all the thermodynamic quantities can be obtained from the 'chopped up' system (a collection of, or rather, an ensemble

---

[12] Consequently, we will not discuss systems interacting with gravitation. Although we say we take the thermodynamic limit, we ignore gravitational interactions among subsystems. For the precise meaning of the technical terms see the entry below of the fourth law of thermodynamics [T2].

of mutually isolated subsystems).[13] This does *not* mean that a macroscopic system actually consists of numerous isolated macroscopic subsystems. It simply means that to obtain thermodynamic observables (macroscopic observables including, e.g., the form factors or correlation functions), any macroscopic system may be regarded equivalent to a collection of macroscopic subsystems obtained by chopping the system up. Since these subsystems are not interacting at all,[14] we may interpret a macroscopic object in equilibrium as a collection of numerous statistically unrelated (i.e., independent)[15] macroscopic objects for the purpose of computing thermodynamic observables. We may express this fact as the partition/rejoining invariance of thermodynamic observables.

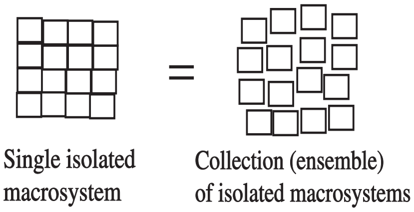


Fig. 1.1.2 An isolated macroscopic system in equilibrium is thermodynamically equivalent to a collection of isolated macroscopic (sub)systems in equilibrium.r

**The relation between thermodynamic limit and finite systems**r
As noted before, the laws of thermodynamics are, precisely speaking, the ones for infinite systems. Then, aren't all the extensive quantities meaningless? That is right; truly meaningful thermodynamic observables for infinite systems are only *densities* of the so-called extensive quantities and their conjugate intensive quantities (called *fields*). That is, the fourth law asserts that proper thermodynamic observables are

---

[13] Needless to say, the fluctuations in small pieces are large. However, if the fluctuations become three orders as large as the original system after chopping it up into one million pieces, the average of all the small piece observables cancel exactly this increase of fluctuations thanks to the law of large numbers (see Sect. 1.2).

[14] If there is an additive conserved quantity (e.g., energy), its total sum must be respected. The reader might think there could be a problem of how to distribute conserved quantities to subsystems, but if the original system is uniform, we can simply distribute them evenly.

[15] **⟪On virtual statistical independence of the parts⟫** It is hard to assert the statistical independence of the resultant pieces. The true meaning of the assertion here is: irrespective of the actual picture, we may assume statistical independence. For a system in equilibrium, we may define temperature. Thermodynamic observables of a system are intact even if the system is in contact with a heat bath at the same temperature. Therefore, even if the pieces obtained by chopping-up a macrosystem are joined after they are in thermal contact with their individual (private) thermal baths, the resultant state must be thermodynamically indistinguishable from the original state of the original macrosystem. Consequently, even if we assume each piece is statistically independent, thermodynamically no difference should be detected.

Furthermore, as stated in the footnote of Borel and chaos (footnote 10), in reality, it is impossible to eliminate external disturbances, so each part would be statistically independent sooner or later.

densities and their conjugate fields. However, in statistical mechanics we cannot handle infinite systems at once. It is meaningful to characterize extensive quantities as quantities asymptotically proportional to the system size (i.e., the amount of materials). Even though the systems we can actually observe are all finite, their sizes are much bigger than microscopic mechanical elements, so the values of thermodynamic observables (densities and fields) are almost indistinguishable from their thermodynamic limit values. Precisely stated, the values of thermodynamic observables for finite systems fluctuate around the thermodynamic limit values with latitudes of $O[1/\sqrt{N}]$ (see Sect. 2.7), where $N$ is the number of microscopic entities in the system. This is the reason why thermodynamics is practical.

**How big is a 'real' finite system?**
Since a real system is finite, we cannot chop it up into infinitely many 'macroscopic' pieces. Then, are we allowed to regard it as an ensemble of sufficiently numerous subsystems? Notice that the 'tiny pieces' obtained by chopping up an ordinary macrosystem into one million pieces are sufficiently macroscopic (we have only to chop up one liter of substance into 1mm cubes; if the substance is a gas at a room temperature at 1 atm, how many molecules are in it?). As we will see in the next section, for the law of large numbers to hold with reasonable accuracy, we do not need such a big number of samples. In short, a realistic macrosystem can be understood as an ensemble of a sufficiently large number of macroscopic subsystems for the law of large numbers to hold.[16]

The partition/rejoining invariance of thermodynamic observables, together with the law of large numbers (see the next section), implies that thermodynamic measurements can be instantaneous, and that their results are reproducible whenever they are measured. These observations are consistent with empirical facts. For example, to measure the compressibility of a gas, no more than 1 ms would be needed. For an identical equilibrium state (up to the observation errors and thermal fluctuations) we can reproducibly repeat measurements with long intervals between them.

Microstates that actually occur during a very short observation time for a macrosystem must be an overwhelming minority among the microstates that can occur under

---

[16] ⟪**How small are molecules?**⟫ Let us have a 'real feel' for how small and consequently how numerous molecules are. The total mass of water (mostly sea water) on the earth is $1.4 \times 10^{21}$ kg $= 7.8 \times 10^{22}$ moles $= 9 \times 10^{22}$ tablespoons (one tablespoon is 15 cm$^3$). The number of water molecules in a tablespoon is five times as large as the number of scoops needed to exhaust sea water. What if we replace the tablespoon with a teaspoon (5 cm$^3$)?

the same equilibrium condition.[17,18] In other words, thermodynamic measurement gives averages over extremely localized microstates sampled from the phase space (sampled from the set of all the microstates). Intermittently repeated measurements must observe quite distinct microstates, so the excellent reproducibility of measurement results implies that almost all locally sampled sets of microstates give the identical thermodynamic result. The most natural and simplest interpretation of these facts is that almost all individual microstates give identical thermodynamic results.

Statistical mechanics assumes the conclusion of the above intuitive considerations. Namely, it assumes that almost all microstates that can occur under the condition that a particular macrostate is observed (we say these microstates are consistent with the macrtostate) individually give the same thermodynamic observable values. Then, thermodynamic quantities should be obtainable from almost any one of the microstates consistent with the equilibrium state of a macroscopic system without any averaging or taking statistics; we do not need statistics, do we? There are many macroobservables for a microstate. How can we extract thermodynamic quantities from all macroobservables? Since thermodynamic observables take identical values for almost all microstates consistent with a given macrostate, it is natural to expect that they may be extracted by averaging evenly over all consistent microstates.

In practice, (as we may see from computer experiments) a system with $10^6$ particles is often large enough to obtain thermodynamic quantities. Therefore, $10^{10}$ 'macroscopic' subsystems can easily be imagined in an ordinary macroscopic object.

---

[17] For example, in classical mechanics during this observation time the system must be able to experience only overwhelmingly small portion of the phase space allowed to the system. Those who have learned classical mechanics must know the *Poincaré time*, which is, roughly speaking, the needed time for a system to itinerate all the representative mechanical states and is known to be much longer than the age of the universe for ordinary macrosystems. Therefore, if the Poincaré time is compared with the observation time, the above assertion about the fraction of the phase space actually observed is quite natural.

[18] ⟪**Boltzmann and Zermelo**⟫ Boltzmann utilized the enormous length of the Poincaré time to show that there was virtually no contradiction between his mechanical derivation of irreversibility and mechanical recurrence, and dismissed Zermelo's logical criticism of 1896. Ernst Zermelo (1871-1953) was in those days an assistant to Planck, who was critical about atomism. Read Section 1.4 of H.-D. Ebbinghaus, *Ernst Zermelo, an approach to his life and work* (Springer, 2008). Zermelo's much more important work is axiomatic set theory, but he never lost interest in statistical mechanics. It was he who translated Gibbs's statistical mechanics book into German (in 1908).

Zermelo's criticism may have given an impression that it is an excessively mathematical criticism without understanding the physical reality of the world, and that Boltzmann seems to have treated it as such. However, as will be discussed later, this criticism seems to have seriously affected Boltzmann and made him realize the true implication of the ergodic hypothesis.

For this reason ensemble theory is applicable to a single macroscopic object.[19] In the actual experiment, it appears we do not need any statistics, but to perform thermodynamic observations on a single macrosystem is in practice to take statistics over its numerous subsystems. This seems to be the most natural interpretation of ensemble theory.[20]

**[T2] The fourth law of thermodynamicsqr** (**r**to Lenard's theorem)
The fourth law of thermodynamics may be stated as:

Thermodynamic observables are either homogeneous functions of degree one or degree zero of the amount of materials in the system.

The name was coined and its significance was stressed by Landsberg.[21] For homogeneous function see the next fine-lettered entry. *Extensive quantities* are the quantities that double when the system size is doubled (while keeping the same thermodynamic state). More precisely, a thermodynamic observable is an extensive quantity, if it is proportional to the amount of materials in the system (if it is a homogeneous function of degree 1 of the amount of materials). Examples are internal energy $E$, volume $V$, magnetization $M$, etc. *Intensive quantities* are those independent of the amount of materials in the system (homogeneous functions of degree zero of the amount of materials). Examples are temperature $T$, pressure $P$, etc. In thermodynamics extensive and intensive quantities appear in pairs as $T$ and $S$, $-P$ and $V$, etc. These pairs are called *conjugate pairs*.[22] Since we must take the thermodynamic limit, all the extensive quantities diverge and become meaningless, so, precisely speaking, we must consider their *densities*. Thus, densities and intensive parameters (*fields*) are the fundamental variables of thermodynamics. When two phases are in equilibrium, field values must be identical, and difference in densities characterizes different phases.

**Homogeneous functions**
If a function of $x_1, x_2, \cdots, x_n$ defined on an appropriate region[23] of $\boldsymbol{R}^n$ ($n$-dimensional real space) satisfies for any positive real number $\lambda$

$$f(\lambda x_1, \lambda x_2, \cdots, \lambda x_n) = \lambda^\alpha f(x_1, x_2, \cdots, x_n), \tag{1.1.1}$$

$f$ is called a *homogeneous function of degree* $\alpha$, where $\alpha$ is a real number.
(1) If $f$ is differentiable, then $\partial f / \partial x_k$ for any $k \in \{1, 2, \cdots, n\}$ is a homogeneous

---

[19] That is, in the thermodynamic limit, if we use the terminologies explained later, a single microscopic state of a macrosystem may be interpreted as a microcanonical ensemble.

[20] ⟪**Khinchin and sum function**⟫**r** It was Khinchin who introduced the concept of sum functions (multivariate functions that can be written as a sum of numerous functions with a few variables) and tried to mathematize the idea here. See A. Ya. Khinchin, *Mathematical Foundations of Statistical Mechanics* translated by G. Gamov (Dover, New York, 1949).

[21] P. T. Landsberg, *Thermodynamics with quantum statistical illustrations* (Wiley, 1961).

[22] ⟪**Conjugate pairs**⟫ They are conjugate pairs with respect to energy: $dE = TdS - PdV + hdM$. We can conceive analogous pairs with respect to entropy: $dS = (1/T)dE + (P/T)dV - (h/T)dM$, so when we mention a conjugate pair and when there is a possibility of confusion, we must state explicitly with respect to what a conjugate pair is defined.

[23] This can be a cone whose apex is at the origin.

function of degree $\alpha - 1$.**r** This can be obtained easily by partially differentiating (1.1.1).
(2) If $f$ is differentiable, a necessary and sufficient condition for $f$ to be a homogeneous function of degree $\alpha$ is (*Euler's theorem* for homogeneous functions)(**q**,**r**to ideal gas, **r**to equipartition)

$$\sum_k x_k \frac{\partial f}{\partial x_k} = \alpha f. \tag{1.1.2}$$

Differentiating (1.1.1) with respect to $\lambda$ and setting $\lambda = 1$, we obtain (1.1.2). To show the converse we need a method to obtain the general solution of a quasilinear partially differential equation (1.1.2).[24] We first make the characteristic equation

$$\frac{dx_1}{x_1} = \frac{dx_2}{x_2} = \cdots = \frac{dx_n}{x_n} = \frac{df}{\alpha f}. \tag{1.1.3}$$

Solving this for general solutions, we obtain

$$x_2/x_1 = \text{constant}, \cdots, x_n/x_1 = \text{constant}, f/x_1^\alpha = \text{constant}. \tag{1.1.4}$$

Therefore, taking an arbitrary differentiable function $h$, the general solution to (1.1.2) reads

$$f = x_1^\alpha h(x_2/x_1, \cdots, x_n/x_1). \tag{1.1.5}$$

This is obviously a homogeneous function of degree $\alpha$.

The *Gibbs relation* reads, for example (the reader must know this relation at least roughly[25]),

$$dE = TdS - PdV + hdM, \tag{1.1.6}$$

where the variables are: $E$ internal energy; $T$ temperature; $S$ entropy; $P$ pressure; $V$ volume; $h$ magnetic field; $M$ magnetization.

$$E = E(S, V, M) \tag{1.1.7}$$

is an extensive quantity = a homogeneous function of degree 1. Therefore, (1) above implies that $T$, $P$ and $h$ are homogeneous functions of degree zero = intensive quantities. (1.1.2) implies

$$E = TS - PV + hM. \tag{1.1.8}$$

Combining this with (1.1.6), we obtain the *Gibbs-Duhem relation*:

$$SdT - VdP + Mdh = 0. \tag{1.1.9}$$

Using the above relations, we can obtain the Gibbs relation for densities: $e = E/V$, $s = S/V$ and $m = M/V$. From (1.1.8) we obtain $e = Ts - P + hm$. On the other

---

[24] As a standard reference see L. C. Evans, *Partial Differential Equations* (AMS 1998).

[25] If not, go to [T5] in Section 1.3 for a brief summary; an introductory material is placed in Supplementary Pages.

hand, from the Gibbs-Duhem relation $sdT - dP + mdh = 0$. Combining $de$ and this, we obtain

$$de = Tds + hdm. \tag{1.1.10}$$

We will encounter the *generalized homogeneous function* **r** in Chapter 5 discussing critical phenomena. It is defined by the following formula instead of (1.1.1)

$$f(\lambda^{\alpha_1} x_1, \lambda^{\alpha_2} x_2, \cdots, \lambda^{\alpha_n} x_n) = \lambda^{\alpha} f(x_1, x_2, \cdots, x_n), \tag{1.1.11}$$

where $\alpha_i$ $(i = 1, \cdots, n)$ and $\alpha$ are real. What is the counterpart of Euler's theorem?

The explanation up to this point can only furnish an intuitive motivation for statistical mechanics, and is not sufficiently elaborate to found statistical mechanics on it. In Sect. 1.3 a more formal introduction of statistical mechanics will be given. The rest of this section is devoted to additional comments. Section 1.2 is a review of rudiments of probability theory and the law of large numbers, because thermodynamic observables can be understood as averages over numerous quantities (that are fluctuating due to the finite size of subsystems).

A quantum mechanics aficionado will immediately object that such chopping up mentioned above is not allowed because it destroys quantum mechanical coherence. However, we are only using empirical facts,[26] which tells us empirically that quantum coherence is not a problem for thermodynamic observables. As empiricists, we must pay due attention to the fact that quantum effects were not recognized by human beings until very recent.

It should be clear by now that 'ergodicity' (read the next fine-lettered entry if the reader feel not very familiar with the concept) has nothing to do with equilibrium statistical mechanics. The traditional (legendary?) viewpoint that ergodicity is crucial is chiefly due to Boltzmann's followers (esp., Ehrenfest) who misunderstood

---

[26] ⟪**Macroscopic quantum coherence**⟫ There are states with global quantum coherence such as in superfluids. In this case the superfluid order may be considered just as magnetic order, so when the system is chopped up or subsystems are joined, the direction of the order (phase) must be respected. The crucial point is that this coherence can be macroscopically controlled (at least in principle); with an appropriate boundary condition the coherence in a given subsystem can be maintained, and when subsystems are joined, the global coherence can be reconstructed. Almost tautologically, not macroscopically observable coherence [more precisely, the main topic of A. J. Leggett, "Macroscopic Quantum Systems and the Quantum Theory of Measurement," Prog. Theor. Phys., Suppl. **69**, 80 (1980)] does not matter thermodynamically. Such coherence can be ignored.

the nature of statistical mechanics.[27] It is still an entrenched view that the statistical mechanical probability is given by the relative sojourn time in a set of microstates, so let us have a critical discussion.

How long does it take to sample a "reasonable" portion of the collection of microscopic states allowed to a macrosystem, or how long does it take for the system to itinerate those microstates? As we have already discussed, it must take eons to sample even a millionth of the phase space.[28] In short, no one can observe the system long enough to substantiate Boltzmann's conception of probability as the proportion of sojourn time for a set of microstates. The important point here is that the idealized thermodynamic observation is an instantaneous observation and has nothing to do with the dynamic behavior of the system during the observation. This implies, as we have already concluded above, that the observation of (almost) any single microstate consistent with a macroscopic system in equilibrium is enough to study the thermodynamics of the system. There is no room for ergodic theory.[29] The problem of ergodicity became regarded fundamental simply because the founding fathers completely misunderstood the origin of statistics.[30]

**†Ergodicity**
Let $\Omega$ be a space (e.g., the phase space) and $T_t$ be a one-parameter family of maps from $\Omega$ into itself (endomorphism) satisfying (i) $T_sT_t = T_{s+t}$ and (ii) $T_0 = 1$ (the

---

[27] For the original meaning of Boltzmann's ergodic hypothesis see Jan von Plato, "Boltzmann's ergodic hypothesis," Archive for History of Exact Sciences **42**, 71 (1991).

[28] ⟪**Short time correlation relevant?**⟫ There is an opinion that the correlation time of the observed data are so short that observed data may be treated almost as independently and identically distributed samples. Consequently, almost instantaneously we can collect samples required by the law of large numbers; the required time is independent of the system size. Therefore, the criticism in the text is irrelevant or not serious. This argument sounds plausible, but we should not forget that the microstates that may be sampled within a short time must lie locally closely to the initial state. The shortness of the correlation time does not logically imply that observed quantities agree globally everywhere in the state space. The other empirical fact that we can reproduce the same result the next day implies that we can obtain the same observed results wherever we start in the state space. That is, the crucial point is that all the microstates give the same observation results, and the shortness of the correlation time is irrelevant.

[29] ⟪**Boltzmann realized...**⟫ Boltzmann noted the discrepancy in time scales pointed out here, and concluded that except for very rare states, all the microstates give the same thermodynamic observables. That is, he reached the same conclusion as ours that is based on thermodynamics and empirical arguments. However, Boltzmann's crucial observation was not understood by his successors, and was soon forgotten. G. Gallavotti, *Statistical Mechanics: a short treatise* (Springer, 1999) Chapter 1 explains Boltzmann's original idea.

[30] Ergodic theory is an extremely important branch of mathematics, so the contribution of the founding fathers of statistical mechanics to motivate the theory is significant. Even a total misconception could produce really deep results if taken seriously. The secret of Magellan was his sincere belief in a wrong map (according to S. Zweik).

identity map). The reader may understand that $T_t(x)$ is the solution of a differential equation at time $t$ with the initial condition $x$ at time $t = 0$. That is, $T_t$ is a time evolution operator. $(\Omega, T_t)$ is called a (continuous time) *dynamical system.*

We can imagine a probability measure $m$[31] on $\Omega$ that is invariant under $T_t$ (i.e., $m(T_t^{-1}(A)) = m(A)$ for any $A \subset \Omega$ for which $m(A)$ is defined). $T_t^{-1}(A)$ is the set of all the points that will visit $A$ in $t$. Thus, $m(T_t^{-1}(A)) = m(A)$ implies that the probability (estimated according to the weight $m$) for an event to occur that guarantees the occurrence of $A$ in $t$ is equal to the probability (estimated according to the weight $m$) that $A$ actually occurs. The reader may understand such $m$ (called an *invariant measure*) as a steady state (a steady state distribution). There is always at least one such measure for any dynamical system with a compact phase space (Krylov-Bogoliubov's theorem).

$(\Omega, T_t, m)$ is called a *measure theoretical dynamical system.* Notice that, generally speaking, a dynamical system $(\Omega, T_t)$ has many (often uncountably many) invariant measures, so there are many measure theoretical dynamical systems $(\Omega, T_t, m)$ for a given dynamical system $(\Omega, T_t)$. Needless to say, the dynamical system does not have any rule to choose a particular one of them.

$(\Omega, T_t, m)$ is called an *ergodic dynamical system*, if for any function $f$ on $\Omega$ that has an expectation value with respect to $m$ the equality

$$\int_\Omega dm(x) f(x) = \lim_{t\to\infty} \frac{1}{t} \int_0^t f((T_s(x_0)) ds \tag{1.1.12}$$

holds for almost all $x_0 \in \Omega$ with respect to $m$ (agreement of the ensemble average and the time average). The invariant measure $m$ satisfying (1.1.12) is called an *ergodic measure.*[32]

The reader must remember that ergodicity is not a property of the dynamical system but of the measure theoretical dynamical system. Therefore, ergodicity of a classical mechanical system does not make any sense until its invariant measure is specified, even if $\Omega$ and the Hamiltonian are chosen. Usually, the phase volume confined to a constant energy 'shell' is chosen as the invariant measure (thanks to Liouville's theorem; we will discuss this in Sect. 1.3). Ergodic theory never answers the question why this invariant measure is the distribution underlying equilibrium statistical mechanics, so even if the corresponding measure-theoretical dynamical system is shown to be ergodic, the conclusion does not justify statistical mechanics. As discussed in the main text, ergodicity is an irrelevant topic for equilibrium statistical mechanics.

Finally, let us briefly reflect on what we mean by mechanics. Stated intuitively, mechanics is a time evolution law that preserves microscopic information. We will

---

[31] An elementary introduction to measure can be found in Y. Oono, *The Nonlinear World* (Springer 2012).

[32] A convenient introduction is P. Walters, *Introduction to Ergodic Theory* (Springer, 1982). Elementary measure theory is a prerequisite. Most physics students will not be able to read the book easily.

later discuss how to quantify information (Sect. 2.4), but what 'microscopic information' is is not often discussed in physics. Here, let us simply understand that it is the information needed to single out a particular microstate, and an important requirement for a mechanics is that it preserves this information. Honestly speaking, it is very likely that we still do not know the true mechanics. The rule of the game of statistical mechanics is to assume that all the microscopic constituents obey a certain mechanics. How can we be sure that all molecules simultaneously obey (some sort of) mechanics in a macroscopic system? No way. This is a metaphysical assumption; metaphysical in the sense that there is no experimental means to check this assertion.[33]

We need a certain metaphysical framework to understand the world[34] just as we need some coordinate system to describe a spatial curve in detail. However, just as what is really there is invariant under the change of coordinate systems, what is real must be invariant under changing metaphysics. Science must be metaphysics-invariant. The implication of this requirement to statistical mechanics does not seem to have been explored.

## 1.2 Law of large numbers

To understand the equilibrium state thermodynamically, we may regard a macroscopic object as a collection of numerous macroscopic objects which look the same macroscopically up to fluctuation, and are statistically independent. That is, thermodynamic observables (densities and fields) may be understood as an average of numerous independently and identically distributed (often abbreviated as *iid*) stochastic variables (for terminologies and notations, see the review beginning on the next page):

$$\frac{1}{n}\sum_{i=1}^{n} X_i, \tag{1.2.1}$$

---

[33] A standard argument against this is that if we do not encounter any difficulty and can explain all what we observe by assuming this 'assumption,' it should be regarded that the assumption has been empirically verified. This sounds very sensible and plausible, but how can we be sure that we have checked all the representative cases? History tells us that it is often the case that we see (observe) only what we wish to see (observe).

[34] E. Schrödinger, *Meine Leben, Meine Weltansicht* (Paul Zsolnay Verlag, 1985).

where $X_i$ is the observable for the $i$-th subsystem. The key mathematics we need is the *law of large numbers*r (rto fluctuation):

> Let $\{X_i\}$ be a collection of iid stochastic variables. For any $\varepsilon > 0$,
>
> $$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - E(X_1)\right| > \varepsilon\right) = 0 \qquad (1.2.2)$$
>
> holds under the condition that the distribution of $X_i$ is not too broad: $E(|X_1|) < \infty$, where $E(Y)$ implies the expectation value of the stochastic variable $Y$.

It is possible to say that this explains why repeated instantaneous thermodynamic observations give exactly the same results again and again.

At this occasion let us review elementary probability theory.[35] What is *probability*? The probability of an event is a measure of the extent of our confidence in the occurrence of the event in 0-1 scale. It behaves like the volume[36] of a geometric object. For example, when the reader draws a card from a stack of playing cards, the certainty of it being a face card ought to be the sum of the certainty for it to be a king and that to be a queen or a jack.

We model an event as a set. First, we consider a set $\Omega$ (called the *whole event*) whose subsets can describe all the possible events under consideration. Its elements are called *elementary events*,[37] and any subset of $\Omega$ is an *event*. Put more plainly, a

---

[35] ⟪**Measure theoretical probability**⟫ The subsequent introduction smoothly flows into the axiomatic formulation of measure theoretical probability due to Kolmogorov. K. Itô, *Introduction to Probability Theory* (Cambridge UP, 1984) may be a good introduction. For those who know rudiments of probability theory, R. Durrett, *Probability: theory and examples* (Wadsworth & Brooks/Cole, 1991) is strongly recommended. It is a beautiful book.

[36] ⟪**Importance of additivity**⟫ Mathematically speaking, it behaves as a ($\sigma$-)additive measure. The additivity is derived in R. T. Cox, "Probability, frequency and reasonable expectation," Am. J. Phys. **14**, 1 (1946) with the aid of a very general axiomatic consideration about 'reasonable expectation.' N. Chater, J. B. Tenenbaum and A. Yuille, "Probabilistic models of cognition: Conceptual foundations," Trends Cognitive Sci., **10**, 287 (2006) contains relevant interesting discussions.

[37] ⟪**Elementary events**⟫ Since events are subsets of $\Omega$, not the elements of $\Omega$ but the singletons (sets such as $\{a\}$ consisting of a single element $a \in \Omega$) should be called elementary events, but it is customary to call elements also elementary events. The elementary event need not be an event that we cannot analyze further, but can be an 'equivalence class' of events obtained when we ignore certain aspects of events we are not interested in. For example, when we cast a dice, we usually ignore where it lands or in which direction its particular face faces. That is, we classify all the events into equivalence classes according to the numbers, and regard these equivalence classes as elementary events. Needless to say, if the reader wishes to use the direction of the faces in a gamble, she must set up elementary events differently.

set $\Omega$ consisting of the totality of elementary events (the events we need not analyze further) is constructed first. The set is called the whole event, any of whose subset is called an event: we say an event occurs if an elementary event belonging to it actually occurs.

> The preceding paragraph may be illustrated with a dice as follows. The elementary events are $e_n$ = "$n$ is rolled" ($n = 1, 2, 3, 4, 5$, or 6), and the whole event is $\Omega = \{e_1, \cdots, e_6\}$. For examples, the event that a prime number is rolled is interpreted as the subset $\{e_2, e_3, e_5\}$, and the event that an even number is rolled corresponds to the subset $\{e_2, e_4, e_6\}$.

If two events $A$ and $B$ do not have any common element ($A \cap B = \emptyset$), we say they are *mutually exclusive*, and these events never occur simultaneously. Thus, it is sensible to assert that if $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B). \tag{1.2.3}$$

This is *additivity*.[38] The equality implies that probability is something like volume or weight of an event (i.e., a set). Something (some elementary event) must happen. Therefore,

$$P(\Omega) = 1, \tag{1.2.4}$$

and it is impossible for nothing to happen:

$$P(\emptyset) = 0. \tag{1.2.5}$$

This concludes the characterization of probability. That is, probability $P$ is an additive set function[39] whose range is $[0, 1]$. $P$ is often called a *probability measure*.

---

[38] ⟪**$\sigma$-additivity**⟫ From (1.2.3) we can conclude that for any finite family of mutually exclusive sets: $\{A_i\}_{i=1}^n$ (here, $A_i \cap A_j = \emptyset$ for any $i, j \in \{1, \cdots, n\}$)

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i).$$

However, we cannot generalize this to an infinite family. Therefore, in the standard axiomatic characterization of probability, the extension of the above formula to $n = \infty$ is required ($\sigma$-*additivity*) instead of (1.2.3).

[39] ⟪**Set function**⟫ A *set function* on a set $\Omega$ is a function that assigns a number to a subset of $\Omega$. To be precise, we may not be able to assign a number to every subset of $\Omega$, so we must specify the family $\mathcal{F}$ of subsets in $\Omega$ to which we may define $f$. However, we will often omit this further specification in this book.

We say event $A$ and event $B$ are *independent*,[40] if

$$P(A \cap B) = P(A)P(B). \tag{1.2.6}$$

Independence implies that knowledge about event $A$'s occurrence or non-occurrence yields no information about event $B$'s occurrence or non-occurrence and vice versa. Do not confuse independence and mutual exclusiveness.

Suppose we know that an event $A \subset \Omega$ has occurred. How should we define the conditional probability $P(B|A)$ for the event that $B$ occurs under the condition that $A$ occurs? Our confidence in simultaneous occurrence of $A$ and $B$ is $P(A \cap B)$. Since we know about $A$, it is plausible that conditional probability should be defined by scaling the probability so that $P(A|A)$ is unity. Therefore, we define the *conditional probability* of $B$ under condition $A$ as

$$P(B|A) = \frac{P(A \cap B)}{P(A)}. \tag{1.2.7}$$

If $A$ and $B$ are statistically independent, note that $P(B|A) = P(B)$. This is intuitively very sensible.[41]

> †**Interpretation of probability**
> There are two major camps as to the interpretation of probability: the *frequentists* and the *subjectivists*.[42] The former asserts, crudely put, that probability of a particular event is the frequency of its occurrences relative to all the number of occurrences of the possible events. The latter points out the difficulty in the interpretation of a unique event (an event that occurs only once, e.g., today's weather), and asserts that some probabilities must be interpreted deductively rather than inductively contrary to the frequentists.[43]

---

[40] **Warning ⟪Pairwise independence is not enough⟫**: Pairwise independence does not imply multiple independence: even if $A, B, C$ are three events that are pairwisely independent, $P(A \cap B \cap C) = P(A)P(B)P(C)$ does not follow. For example, suppose we throw a fair coin twice, Let $A$ be the event that the first result is head, $B$ the event that the second result is head, and $C$ the event that the first and the second results agree (both head or both tail). The event $A \cap B \cap C$ implies that head appears twice, so $P(A \cap B \cap C) = 1/4$, disagreeing with $P(A)P(B)P(C)$. If we have many events, to specify statistical independence of all the events, we must specify independence in detail.

[41] According to a remark attributed to Mark Kac, probability theory is "Measure theory with a Soul."

[42] D. Gillis, *Philosophical theories of probability* (Routledge, 2000) has a summary. However, the point of view advocated here is not discussed at all.

[43] Propensity interpretation by Popper, interpretation in terms of degree of confirmation by Carnap, etc. See Gillis cited already. We may say that these are products of humanistic fallacy (= fallacy common to the conventional humanities) that forgets our being living organisms.

> The characterization of probability in the text is as a measure of confidence, so the reader might think that the author is on the subjectivist side. Despite characterizing probability as a measure of confidence, he believes that this subjective measure *must* agree at least approximately with our experiences. Measure of confidence must agree with relative frequency with which the event occurs, when frequency interpretation is possible. The author's assertion is that probability is a measure of confidence inductively (empirically) supported by relative frequency. 'Non-frequentists' forget an important fact that our mental faculty is a product of evolution; our confidence may not be tied to our own experience, but is tied to the result of our phylogenetic learning = the totality of experiences by our ancestors during the past 4 billion years.[44]
>
> It was Laplace who compiled classic probability theory (of Bernoulli, Fermat, de Morgan and others).[45] His basic idea was: the relative number of equally probable elementary events in a given event is the probability of the event. Here, 'equally probable' implies that there is no reason for any difference.[46] This is exactly the same as Boltzmann's interpretation of probability. If we feel that there is no reason to detect any difference between event A and event B, then we feel A and B are equally likely. This feeling is often in accordance with our experience thanks to the evolution process that has molded us. That is, our feeling and the 'logic' behind Laplace's and Boltzmann's fundamental ideas have the basis supported by phylogenetic learning; Probability is always (disguised) relative frequency, and the subjective interpretation is a mere figment of the humanistic fallacy that ignores the fundamental fact that we are living organisms.

The whole event $\Omega$ and a probability measure[47] $P$ on it define a *probability space* $(\Omega, P)$.[48] A *stochastic variable*$_{\mathbf{r}}$ $X$ is a function defined on a probability space:

---

[44] ⟪**Our innate statistics capability**⟫ Interestingly, it may well be that we have a mental device to take statistics unconsciously. Not only us but as the following paper reviews a variety of animals in different ecological contexts behave in manners consistent with predictions of Bayesian updating models: T. J. Valone, "Are animals capable of Bayesian updating? An empirical review," Oikos, **112**, 252 (2006). The following paper may also be related: P. J. Bayley, J. C. Frascino and L. R. Squire, "Robust habit learning in the absence of awareness and independent of the medial temporal lobe," Nature **436**, 550 (2005). A slow learning process with trial and error can occur unconsciously.

[45] ⟪**Laplace's demon**⟫ The Laplace demon that can predict everything deterministically was introduced to illustrate its absurdity and to promote probabilistic ideas in the introduction to his *Essai philosophique sur les probabilitès*. [1814, this year Napoleon abdicated; Fraunhofer lines; Berzelius introduced the modern chemical symbols now in use.]

[46] **Warning** ⟪**Ignorance and probability**⟫ Do not confuse this idea with the assignment of equal probabilities due to ignorance (= if we do not know anything about events $A$ and $B$, we set $P(A) = P(B)$). The latter is a mere abuse of probability without any justification. Appreciate that Laplace's idea is diametrically different. He says that the same probabilities are assigned if no difference can be found even after scrutiny.

[47] The reader can intuitively understand that it is a sort of volume.

[48] ⟪**Probability space**⟫ Mathematically, we must also specify what events are allowed to have probabilities. This is defined by specifying a family $\mathcal{F}$ of measurable subsets of $\Omega$. Therefore,

$X = X(\omega)$ $(\omega \in \Omega)$ (its value may be a vector, a complex number, etc.).[49] The *expectation value* (mean or average) of a stochastic variable $X$ is defined by[50]

$$E(X) = \int X(\omega)dP(\omega). \tag{1.2.8}$$

The *indicator* $\chi_A$ of an event $A$ is a function satisfying the following conditions:

$$\chi_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A, \text{ that is, if event } A \text{ occurs,} \\ 0, & \text{if } \omega \notin A, \text{ that is, if event } A \text{ does not occur.} \end{cases} \tag{1.2.9}$$

Its expectation value gives the probability of the event $A$:

$$P(A) = E(\chi_A). \tag{1.2.10}$$

*Conditional expectation value* may be defined analogously. Under the condition that event $A$ occurs, the conditional expectation value $E(X|A)$ of $X$ is defined as

$$E(X|A) = \sum_{\omega_i \in \Omega} X(\omega_i)P(\omega_i|A) \tag{1.2.11}$$

in terms of conditional probabilities.[51] $E(X|A)$ may be understood as a stochastic variable dependent on $A$, so its average over $A$ is meaningful. In particular,**r** if $A_i \cap A_j = \emptyset$ $(i \neq j)$ and $\cup_i A_i = \Omega$ (i.e., $\{A_i\}$ is a *partition* of $\Omega$), then the average over $\{A_i\}$ of $E(X \mid A_i)$ is the unconditional average:

$$E(E(X|A_i)) = E(X). \tag{1.2.12}$$

---

mathematicians write the probability space as $(\Omega, \mathcal{F}, P)$. However, throughout this book it is avoided to be conscious about the axiom of choice and the like.

[49] A stochastic variable is not simply defined as a function on $\Omega$ but that on the probability space, because $X$ as a function on $\Omega$ must be measurable with respect to $P$, but in this book we will not mention such things.

[50] This integral is a Lebesgue-Stieltjes integral, but the reader can intuitively understand it just as the usual Riemann integral as a sum of the product of the representative value $X(\omega)$ of $X$ on the infinitesimal set $d\omega$ and the probability $dP(w) \equiv P(d\omega)$ for the event $d\omega$ to occur. If $\Omega$ is discrete, the integral becomes the sum $E(X) = \sum_i X_i P_i$, where $X_i$ is the value of $X$ for event $i$ and $P_i$ is the probability of this event.

[51] If the probability space is specified with a continuos parameter, we must define conditional probability to the conditions with zero probability, so we need a much more careful definition of conditional quantities. This book does not require such cases, so we do not discuss probability zero conditions.

This is an obvious equation, but is important in mean field theory (Sect. 5.8).

The *variance* of $X$ is defined as

$$V(X) = E([X - E(X)]^2) = E(X^2) - E(X)^2. \tag{1.2.13}$$

Its square root $\sigma(X) = \sqrt{V(X)}$ is called the *standard deviation* of $X$. For two stochastic variables $X$ and $Y$

$$C(X, Y) = E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y) \tag{1.2.14}$$

is called the *correlation* between $X$ and $Y$, which shows up often when we wish to study fluctuations (Sections 2.7 and 2.8).

Statistical independence of two stochastic variables$_{\mathbf{r}}$ $X_1$ and $X_2$ is defined by

$$E(f(X_1)g(X_2)) = E(f(X_1))E(g(X_2)) \tag{1.2.15}$$

for any functions[52] $f$ and $g$ of the stochastic variable. Think of the relation between this independence and the independence of events above. The correlation between $X_1$ and $X_2$ vanishes if they are statistically independent, but vanishing of correlations does not imply independence of stochastic variables.

Now, we can demonstrate the (weak[53]) law of large numbers (1.2.2).[54] First, let us show this 'honestly,' and then 'rederive' it in a theoretical physicist's fashion.

The key to an honest proof is *Chebyshev's inequality*

$$a^2 P(|X - E(X)| \geq a) \leq V(X). \tag{1.2.16}$$

This can be shown as follows (let us redefine $X$ by shifting as $X - E(X)$ to get rid of $E(X)$ from the calculation):

$$V(X) = \int X^2 dP(\omega) \geq \int_{|X| \geq a} X^2 dP(\omega) \geq a^2 \int_{|X| \geq a} dP(\omega) = a^2 P(|X| \geq a). \tag{1.2.17}$$

---

[52] 'Any functions' here means 'any integrable functions.' Such statements will not be explicitly mentioned that are obvious to mathematicians but 'nuisance' to most physicists.

[53] As to the strong law of large numbers, see Supplementary Pages.

[54] In the following the assertion is proved under a stronger condition that $V(X)$ is finite. To prove the law under the condition $E(|X_1|) < \infty$ requires some tricks.

We wish to apply Chebyshev's inequality to the sample average $(1/n)\sum X_k$. Its expectation value is $E(X_1)$ (here $X_1$ is used as a representative variable, but all the variables give the identical result) and its variance is computed as

$$V\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}V\left(\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} V(X_k) = \frac{1}{n}V(X_1) \tag{1.2.18}$$

thanks to the independence of each variable (see (1.2.15)). From this immediately follows

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - E(X_1)\right| \geq \varepsilon\right) \leq \frac{V(X_1)}{\varepsilon^2 n}. \tag{1.2.19}$$

For our purpose $n \gg 10^{10}$, so $\varepsilon$ can be taken much smaller than $10^{-5}$. Thus, we almost exclusively encounter the microstates that are extremely close to the average behavior. If we adopt the intuitive understanding of the ensemble discussed in the preceding section, the reproducibility of thermodynamic observables is due to the law of large numbers. This is the significance of the law to statistical mechanics.

**How to use Chebyshev's inequality**
(1.2.19) may be used as follows. Suppose we wish to estimate the empirical expectation value within the error of $10^{-2}$. We tolerate larger errors if they occur only rarely, say, less than once in 1000 such calculations. What $n$ should we choose?[55] Obviously, we must require

$$\frac{V(X_1)}{10^{-4}n} \leq 10^{-3} \Rightarrow n \geq V(X_1)10^7. \tag{1.2.20}$$

If the variance is of order 1, we should use at least about $10^7$ samples. In the context of statistical mechanics this is nothing.
(1) [High dimensional numerical integration]**q**
Let us consider the problem of numerically evaluating a high-dimensional integral:

$$I = \int_0^1 dx_1 \cdots \int_0^1 dx_{1000} f(x_1, \cdots, x_{1000}). \tag{1.2.21}$$

If we wish to sample (only) two values for each variable, we need to evaluate the function at $2^{1000} \sim 10^{300}$ points. Such sampling is of course impossible. This integral can be interpreted as the average of $f$ over the 1000 dimensional unit cube:

$$I = \frac{\int_0^1 dx_1 \cdots \int_0^1 dx_{1000} f(x_1, \cdots, x_{1000})}{\int_0^1 dx_1 \cdots \int_0^1 dx_{1000}} \tag{1.2.22}$$

---

[55] Here, 'once' means a single set of $n$ consecutive observations or experiments to compute one empirical expectation value. We do not wish to have more than one bad empirical expectation value obtained in this manner out of one thousand empirical expectation values.

Therefore, randomly sampling the points $\boldsymbol{r}_k$ in the cube, we can obtain

$$I = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\boldsymbol{r}_k). \tag{1.2.23}$$

How many points should we sample to estimate the integral within $10^{-2}$ error, if we allow larger errors at most once out of 1000 such calculations? We already know the answer from (1.2.20): $V(f(X_1))10^7$. The variance of the value of $f$ is of order $\max |f|^2$, if $f$ is bounded, so we may scale it to be of order unity. Compare this number with $10^{300}$ above and appreciate the power of randomness.[56] This is the principle of the Monte Carlo integration. Notice that the computational cost does not depend on the dimension of the integral.

(2) [Law of large numbers for empirical probability]
How many times should we throw a coin to check its fairness? The empirical probability for Head is given by $N_H/N$, where $N$ is the total number of trials and $N_H$ the number of trials resulting in Head. The expectation value of $N_H/N$ is the probability of Head $p_H$. Let $X_i$ be the indicator of the Head event for the $i$-th trial. Its expectation value is also $p_H$ and $N_H = \sum_i X_i$. Let $V (\leq 1/4)$ be its variance. Then, the Chebyshev inequality implies

$$P\left(|N_H/N - p_H| \geq \varepsilon\right) \leq \frac{V}{\varepsilon^2 N}. \tag{1.2.24}$$

Therefore, the more unfair the easier to find the fact (express $V$ in terms of $p_H$), but, for example, 10% unfairness is not very easy to detect.

Next, let us compute the density distribution function (for the terminology, see the fine-lettered entry below (1.2.28)) $f_n$ of the $n$ sample average by a formal calculation with the aid of a *generating function* technique$_{\mathbf{q}}$ (can the reader tell where the following proof is incomplete?):

$$f_n(x) = E\left[\delta\left(\frac{1}{n}\sum_{i=1}^{n} X_i - x\right)\right]. \tag{1.2.25}$$

Its *characteristic function*$_{\mathbf{q}}$ $g_n(\xi)$[57] may be given by

$$g_n(\xi) \equiv \int dx\, e^{i\xi x} f_n(x) = E\left(\prod_{i=1}^{n} e^{i\xi X_i/n}\right) = E(e^{i\xi X_1/n})^n. \tag{1.2.26}$$

---

[56] **⟪Constructive role of randomness⟫** Read the article by N. Williams, “Mendel’s demon” in Current Biology **11**, R80-R81 (2001) and discuss the positive meaning of randomness. This is a preview of M. Ridley, *The Cooperative Gene, How Mendel’s Demon explains the evolution of complex beings* (The Free Press, 2001).

[57] We freely exchange the order of integrations. All of them can be justified a la Fubini (sometimes with some additional technical conditions).

Notice that $X_i$ are independent. Let us write the characteristic function of $X_1$ as $\omega(\xi)$. Then, we have

$$g_n(\xi) = \omega(\xi/n)^n. \tag{1.2.27}$$

For very large $n$, the variable in $\omega$ is very close to zero for any fixed $\xi$. Since the distribution has a finite mean $m$ and variance, $\omega$ must be twice differentiable.[58] We may write $\omega(\xi/n) = 1+im\xi/n+O(1/n^2)$. Putting this into (1.2.27), we obtain

$$\lim_{n\to\infty} g_n(\xi) = e^{im\xi}. \tag{1.2.28}$$

This implies that the density distribution of $(1/n)\sum X_i$ (weakly[59]) converges to $\delta(x-m)$. This is exactly what the reader should have expected.[60]

> **Distribution function and density distribution function**
> Let $X(\omega)$ be a real stochastic variable. The probability $F(x)$ for $X \le x$ is called the *distribution function* of $X$:
>
> $$F(x) = P\left(\{\omega \,|\, X(\omega) \le x, \omega \in \Omega\}\right). \tag{1.2.29}$$
>
> Here, the indicator of $X \le x$ may be written in terms of the unit step function $\Theta$[61] as $\Theta(x - X(\omega))$, so we have
>
> $$F(x) = \langle \Theta(x - X(\omega))\rangle. \tag{1.2.30}$$

---

[58] ⟪**Bochner's theorem**⟫ If the distribution has the $m$-th moment, then its characteristic function is $m$-times differentiable. A characterization of the characteristic function is given by *Bochner's theorem*, which is worth memorizing:
Theorem [Bochner] A necessary and sufficient condition for $g(\omega)$ to be the characteristic function of a distribution is:
(i) $g(0) = 1$.
(ii) $g$ is continuous.
(iii) Positive definite: let us choose an arbitrary positive integer $p$, arbitrary reals $x_i$ and arbitrary complex numbers, $\alpha_i$ $(i = 1, \cdots, n)$. Then,

$$\sum_{i,j} \overline{\alpha_i}\alpha_j g(x_i - x_j) > 0.$$

[59] Weak convergence $f_n \to f$ means $\int f_n\varphi\,dx \to \int f\varphi\,dx$ for 'any' $\varphi$ taken from an appropriate function set. In the present case we may take the set of smooth functions with bounded support, that is, the functions that vanish outside some interval.

[60] The 'proof' of the law of large numbers above is not rigorous, because the integration range of (1.2.26) is usually not finite. The reason why the above argument is often justifiable is that in many cases the contribution of $f_n$ for large $|x|$ is small.

[61]
$$\Theta(x) = \begin{cases} 1, & \text{for } x \ge 0, \\ 0, & \text{for } x < 0. \end{cases}$$

The derivative $f(x)$ of $F$ is called the *density distribution function* of $X$. Physicists often call this the distribution function.[62] Since the derivative of the unit step function is the delta function $\delta(x)$, we can write

$$f(x) = \langle \delta(x - X(\omega)) \rangle. \tag{1.2.31}$$

This was our starting point.

**†Kolmogorov's 0-1 law**
Ideally, a macroscopic system (or an ensemble) contains infinitely many statistically independent subsystems. If an observation result of the ensemble is not affected by a finite number of samples, then the probability that observation result becomes equal to a particular value $c$ is 0 or 1. This is Kolmogorov's 0-1 law. This is more basic than the law of large numbers.[63] In the present context, this guarantees the reproducibility of the thermodynamic observable values without error.

For example, for an iid stochastic variable sequence $\{X_i\}_{i=1}^{\infty}$ let us consider an empirical expectation value

$$\frac{1}{n}(X_1 + \cdots + X_n). \tag{1.2.32}$$

If this converges, the result does not depend on a finite number $m$ of observables $X_{i_1}, \cdots, X_{i_m}$, so the event that the empirical expectation value converges to $c$ in the $n \to \infty$ limit occurs with probability 0 or 1 for any $c$. That is, if the expectation value exists, then it is definite (also if it diverges, its divergence is definite).

The law of large numbers and its refinements are the pillars supporting statistical mechanics.**r** One refinement is the central limit theorem which is closely related to renormalization group theory that we will discuss in Chapter 5. Another refinement is *large deviation theory* (LD theory) which is related to the thermodynamic fluctuation theory (Sect. 2.7, 2.8).

When we obtain the law of large numbers, we divide the empirical sum $\sum_{i=1}^{n} X_i$ with $n$ (we assume $X_i$ to be iid stochastic variables as above). Consequently, we cannot detect deviations of this sum from its representative behavior. However, without this division the sum diverges and cannot tell us anything useful. Can we make a quantity that obeys a distribution that does not spread all over nor shrink to a $\delta$-function by dividing the empirical sum with some factor that is smaller than $n$ but still increasing with $n$? This is the basic motivation for the central limit theorem. As can be seen from Chebyshev's inequality, the spread of the empirical sum is $O[\sqrt{n}]$. Therefore, let us divide the empirical sum with $\sqrt{n}$ instead of $n$. Let

---

[62] Therefore, to distinguish it from the more standard usage of the terminology $F$ is sometimes called the *cumulative distribution function.*

[63] For more details, see Section 1.7 of R. Durrett, *Probability: theory and examples* (Wadsworth & Brooks/Cole, 1991).

us compute the characteristic function of

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i. \tag{1.2.33}$$

For simplicity, we assume that the expectation value of $X_i$ is zero, and its variance is unity. Furthermore, we assume that $E(|X_1|^3) < \infty$. The characteristic function $g_n$ of the density distribution function $f_n(y)$ of $Y_n$ may be obtained by mimicking the 'theoretical physics proof' of the law of large numbers as

$$g_n(\xi) = \omega(\xi/\sqrt{n})^n. \tag{1.2.34}$$

Since we have assumed that the third moment of $X_i$ exists, $\omega$ is three times differentiable, so we may write

$$\omega(\xi/\sqrt{n}) = 1 - \xi^2/2n + O(1/n^{3/2}). \tag{1.2.35}$$

Therefore, we obtain

$$\lim_{n\to\infty} g_n(\xi) = e^{-\xi^2/2}, \tag{1.2.36}$$

which is the characteristic function of a Gaussian distribution with average 0 and variance 1. Thus, in the general case of $X_i$ with expectation $m$ and variance $V$, asymptotically in the $n \to \infty$ limit

$$P\left( \frac{1}{\sqrt{nV}} \sum_{i=1}^{n} (X_i - m) \le y \right) \to \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \tag{1.2.37}$$

holds, which is called the *central limit theorem*.[64] This central limit theorem shows us small fluctuations that cannot be detected if we normalize the observables with the system size.

If the variables are not iid, but strongly correlated with each other, what happens to the central limit theorem? We will learn in Chapter 5 that renormalization group theory is a theory attempting to answer this question. The central limit theorem does not appear in elementary statistical mechanics except for the problems such as the following.

**Elementary application of central limit theorem**$_{\mathbf{q}}$

1 mole of argon gas is in a cubic box of volume 20 liters. This cube is halved with

---

[64] **⟪Central limit theorem references⟫** The bible of the central limit theorem is W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II (Wiley, 1971) Chapters III and XVII. Chapter 2 of R. Durrett, *Probability: theory and examples* (Wadsworth & Brooks/Cole, 1991) may be more accessible.

a membrane. What is the probability $P$ for one half of the cube to have $2 \times 10^{-3}\%$ atoms more than the other or more? Estimate the order of $\log_{10} P$. We assume that atoms are classical distinguishable particles.

Let $X_i$ be a stochastic variable such that it is 1 if the $i$-th particle is in the right half, and 0, otherwise. Its mean is 1/2 and the variance is 1/4. The event whose probability we wish to know is characterized by

$$\sum_{i=1}^{N_A}(X_i - 1/2) \geq 10^{-5} \times \frac{N}{2}. \tag{1.2.38}$$

In this question $N_A$ is equal to the Avogadro constant $N_A = 6.02 \times 10^{23}$.

$$P = P\left(\sum_{i=1}^{N_A}(X_i - 1/2) \geq 10^{-5}\frac{N_A}{2}\right) = P\left(\frac{1}{\sqrt{N_A/4}}\sum_{i=1}^{N_A}(X_i - 1/2) \geq 10^{-5}\sqrt{N_A}\right). \tag{1.2.39}$$

Its right hand side may be estimated approximately with the aid of the central limit theorem. Let us write $\varepsilon = 10^{-5}\sqrt{N_A}$. Then,

$$P = \frac{1}{\sqrt{2\pi}}\int_{\varepsilon}^{\infty} e^{-x^2/2}dx = -\frac{1}{\sqrt{2\pi}}\int_{\varepsilon}^{\infty}\frac{1}{x}\frac{d}{dx}e^{-x^2/2}dx \simeq \frac{1}{\sqrt{2\pi}\varepsilon}e^{-\varepsilon^2/2}. \tag{1.2.40}$$

That is,

$$P(2 \times 10^{-3}\% \text{ imbalance}) \sim \frac{1}{\sqrt{2\pi N_A}}10^5 e^{-N_A \times 10^{-10}/2}. \tag{1.2.41}$$

As can be seen from $\log_{10} P \sim -1.3 \times 10^{13}$, the probability is overwhelmingly small. (However, this estimate is very crude, we should not be surprised that the answer is inaccurate by one or 2 orders.) □

The law of large numbers tells us that for any $\varepsilon > 0$ $P(|\sum X_i/n - E(X)| > \varepsilon)$ decays to zero as $n \to \infty$.r One possible refinement is to characterize this decay rate. Generally, this decay is exponential. We ask a bit more precise question: How does

$$P\left(\frac{1}{n}\sum X_i \sim y\right) \tag{1.2.42}$$

decay (if it decays at all) as $n$ is increased, where $\sim y$ roughly means that the value is in the small range around $y$? It is known that if $\{X_i\}$ are not very strongly dependent on each other and not too broadly distributed, then there is a function $I(y)$ called the *rate function* or *large deviation function* such that[65]

$$P\left(\frac{1}{n}\sum X_i \sim y\right) \approx e^{-nI(y)}. \tag{1.2.43}$$

Here, $\approx$ means that the logarithmic ratio of both sides is asymptotically unity. The rate function$_{\mathbf{r}}$ is a nonnegative convex function with unique zero: $I(E(X)) = 0$.[66] That is, $I(y) > 0$ unless $y = E(X)$. Thus, this gives us a *variational principle* for the observable value: "$I(y)$ is minimized at $y = E(X)$." Often we can expect close relations between variational principles and the law of large numbers. In our context this variational principle is actually the Gibbs variational principle and the large deviation theory is the thermodynamic fluctuation theory that we will study in Sect. 2.7 and 2.8.

## 1.3 Setting up statistical mechanics

In Section 1.1 we considered the relation between statistics and macroscopic observables, and concluded that if we wish to have only equilibrium thermodynamics, we may regard a macrosystem under study as a collection of statistically independent macroscopic subsystems. Thus, we reached a fundamental idea that thermodynamic quantities can be collected by averaging observables over all the microstates allowed to the macroscopic subsystems. However, this is only heuristics; no one has been able to justify the idea really mathematically, so we construct statistical mechanics on some basic principles.

**The reader can start to read this book from this section.**$_{\mathbf{r}}$

The aim of statistical mechanics is to relate microstates of a system described by mechanics and macrostates described by thermodynamics. To this end we need a correspondence (or translation) between mechanical observables and thermodynamic

---

[65] ⟪**Large deviation references**⟫ There seems to be no introductory textbook kind enough to physics students. If the reader wishes to have an intuitive feeling, the introduction and Appendix to Y. Oono, "Large deviation and statistical physics," Prog. Theor. Phys., Suppl. **99**, 165-205 (1989) may be of some use. J.-D. Deuschel and D. W. Stroock, *Large Deviation* (Academic Press, 2000) is a standard textbook. See also H. Touchette, "The large deviation approach to statistical mechanics," Phys. Rep. 478, 1-69 (2009).

[66] Some technical conditions are also assumed for $I$.

observables. Also we must know the statistics of microstates consistent with a given macrostate. That is, we need a *translation rule* and a *statistical principle.*

A macrostate is a state uniquely specified by equilibrium thermodynamics. It is specified by internal energy and a few work coordinates, that is, by a set of the thermodynamic coordinates$_{\mathbf{r}}$ of the system (it is reviewed below with fine letters).[67] A microstate is a state regarded as elementary by (classical or quantum) mechanics. That is, classically, a microstate is a point in the phase space. Quantum mechanically, a microstate is a one-dimensional subspace of the vector space of all the allowed states. However, a microstate in this sense can be an extremely strange state such as a superposition of energy eigenstates with quite disparate energies. Therefore, in this book, quantum mechanically, *only the eigenstates of the Hamiltonian describing the system are regarded microstates.*

Statistical mechanics assumes that there are many microstates that are consistent with a macrostate. The first step to formulate statistical mechanics must be to specify the set of microstates consistent with a given macrostate. To this end we must describe the thermodynamic coordinates in terms of mechanical observables. Since thermodynamic quantities and mechanical quantities are not the same, we need a translation or correspondence between them. To identify mechanical energy and thermodynamic internal energy should be without any controversy.[68] Work coordinates are such as volume and magnetization and are with mechanically clear meanings and are microscopically unambiguous, so the correspondence of work coordinates and corresponding microscopic mechanical observables should be without

---

[67] **⟪Significance of thermodynamic coordinates⟫** (**r**to [T5], **r**to Sect. 2.6) The vital importance of thermodynamic coordinates must be emphasized here. The following two points are crucial if one wishes to understand thermodynamics:

1) Thermodynamic coordinates consist of internal energy $E$ and work coordinates $\{X_i\}$ that are described in terms of (macroscopic) mechanics and electromagnetism;

2) A macrostate is an equivalence class of macroscopic states sharing the same thermodynamic coordinate values.

A macroscopic state of macroscopic object cannot be specified by the thermodynamic coordinates alone, because there are many other macroscopically observable properties such as locations, colors, etc. of the systems. Thus, precisely speaking, a macrostate in thermodynamics (a thermodynamic state) is an equivalence class of equilibrium states according to the thermodynamic coordinate values (disregarding other macroscopic features). To describe entropy $S$ as a function of thermodynamic coordinates implies to give the most detailed description of a macrostate possible within the framework of thermodynamics. Therefore, if statistical mechanics determines a functional relation $S = S(E, \{X_i\})$, it implies that statistical mechanics perfectly reproduces thermodynamics.

[68] It is stated 'without controversy' here, but this was not as obvious as we believe now. B. Widom writes as follows: "Is that obvious? If it now seems obvious it is only because we have given the same name, energy, to both the thermodynamic and the mechanical quantities, but historically they came to be called by the same name only after much experimentation and speculation led to the realization that they are the same thing." (B. Widom, *Statistical Mechanics, a concise introduction for chemists* (Cambridge UP, 2002))

any problem. Consequently, the set of microstates $\omega(E, \{X_i\})$ (henceforth, the set of needed work coordinates will be denoted collectively as $\{X_i\}$) corresponding to (consistent with) the thermodynamic state (macrostate) specified by the thermodynamic coordinates $(E, \{X_i\})$ is classical-mechanically defined as the subset of the phase space for which the mechanical energy agrees with $E$ and the work coordinates (collectively denoted) with $\{X_i\}$. Quantum-mechanically, $\omega(E, \{X_i\})$ is a vector space spanned by the eigenkets $|\ \rangle$ belonging to the eigenvalue $E$ of the system Hamiltonian that satisfy $\langle\ |\hat{X}_i|\ \rangle = X_i$, where $\hat{X}_i$ is the operator (observable) corresponding to the work coordinate $X_i$.

Here, since observable values are always inevitably with some error, the correspondences must allow certain ‘leeway,’ that is, the agreements between the microscopic values and the corresponding macroscopic values are allowed to have certain (small) errors. Therefore, precisely speaking, $\omega(E, \{X_i\})$ is defined as the totality of microstates specified by the thermodynamic coordinates $(E', \{X_i'\})$ satisfying $E - \delta E < E' \le E$, $X_1 - \delta X_1 < X_1' \le X_1$, etc.

**[T3] “Thermally Isolated system” and internal energy**r

In the text above, we have assumed that the system is isolated for simplicity. In this case, the system mechanical energy is the internal energy of the system. However, as already noted, many thermodynamic states cannot be realized in strictly isolated systems as exemplified by a non-zero magnetization state of a ferromagnet above its critical temperature. Imposing a conjugate field of a work coordinate is a mechanical operation. Let us call an isolated system with imposed conjugate fields a *thermally isolated system*.[69]

A system isolated thermally can be described by a Hamiltonian of the following structure:r

$$H = H_0 - h_i \hat{X}_i. \tag{1.3.1}$$

where $H_0$ is the Hamiltonian of the system in isolation, $\hat{X}_i$ is the operator corresponding to the work coordinate $X_i$, and $h_i$ is its conjugate field.

Thermodynamically, since the system is imposed the conjugate field $h_i$ the natural thermodynamic potential is $\eta = E - h_i X_i$, which corresponds to enthalpy. That is, the expectation value of $H$ is not the internal energy but the (generalized) enthalpy. However, since the expectation value of $\hat{X}_i$ is the work coordinate, the internal energy of the microstate is obtainable as $E = \langle|H|\rangle + h_i\langle|\hat{X}_i|\rangle$. Thus, even in thermal isolation under the influence of conjugate fields the thermodynamic coordinate values of each microstate is definite.

---

[69] Thermal isolation is not the same as adiabaticity. In thermodynamics, an adiabatic process is a process without any net heat exchange; for example a system attached to a heat bath but if there is no exchange of heat between the system and the bath, any process in the system is an adiabatic process. Notice that the system cannot be described as a pure mechanical system. In contrast, under thermal isolation the system is described as a purely mechanical system.

Incidentally, the word ‘adiabatic’ also appears in quantum mechanics. It means a very slow process without causing any level crossing. In this book this process is called *quantum-mechanical adiabatic process.*

Notice that the internal energy is the expectation value of $H_0$, but it is not necessarily its eigenvalue; we may interpret $h_i\hat{X}_i$ as the potential energy of the system in the external field $h_i$, but we should not forget that this field modifies the system itself microscopically.

**[T4] Simple systems and composite systems r**
In thermodynamics, it is not rare that we must handle a system that may be understood as an interacting collection of simpler systems (subsystems) such as a box divided by a movable wall (piston). In such a system, even if the system is made of a single substance, it is only piecewisely uniform and some of fields (intensive variables) assume different values in different subsystems. Since statistical mechanics should reproduce thermodynamics, such composite systems must also be taken care of by statistical mechanics. However, a wise way to utilize statistical mechanics is to leave what thermodynamics can do to thermodynamics. Thus, in this book, we consider only the statistical mechanics of simple systems. A clearer statement follows.

A *simple system* is a system that exhibits no spatial inhomogeneity in equilibrium due to some constraints not intrinsic to the system (say, by walls or by external fields). In other words, a system is called a simple system which never exhibits any spatial nonuniform of intensive variables (fields) in equilibrium. Spatial nonuniformity due to phase coexistence may occur in a simple system. In this book we discuss only the statistical mechanics of simple systems,[70] and composite systems are left to thermodynamic considerations.

**[T5] The first and the second law of thermodynamics[71] r**
Here, we confine our discussion to a *closed system* that does not exchange matter with its environment. Furthermore, we do not consider chemical reactions. We will return to these topics in Sect. 2.9 devoted to elementary chemical reactions. In the following the term 'adiabatic' may be understood naively for the time being, but see the footnote 77.

The *first law of thermodynamics* tells us:
(i) The equilibrium state of an isolated system is uniquely specified by a set of extensive variables including internal energy r: $(E, \{X_i\})$.[72,73] Here,

---

[70] However, systems under uniform gravity or uniformly rotating systems may be considered occasionally.

[71] Systematic exposition such as the proof of the existence of entropy is not given here. It is explained in an introductory material posted in the Supplementary Pages.

[72] ⟪**Thermodynamic state**⟫ In the text, it is written, following the traditional explanation, that in equilibrium the description of a macrosystem becomes simple in terms of a few thermodynamic variables. However, as already noted in footnote 67, even in equilibrium, a macroscopic system cannot be described completely macroscopically in terms of thermodynamic variables alone. Thermodynamics does not actually claim that in equilibrium a macroscopic system can be described in terms of the thermodynamic coordinates alone. What thermodynamics asserts is that if we introduce an equivalence class of equilibrium states with respect to the thermodynamic coordinates, then a complete theoretical system (called thermodynamics) is possible for the set of these states. These equivalence classes are called (equilibrium) thermodynamic states.

[73] ⟪**Thermodynamic space**⟫ The space spanned by the thermodynamic coordinates is called

> $X_i$ are chosen to be *work coordinates* that are described by mechanics (including electrodynamics).
>
> (ii) There is a process[74] called a *quasistatic reversible adiabatic process* such that $dE = \sum_i f_i dX_i$, where $f_i dX_i$ is the work done to the system when the work coordinate $X_i$ is changed infinitesimally and $f_i$ must be a function of $(E, \{X_i\})$.[75]

Whether $(E, \{X_i\})$ gives a respectable thermodynamic description of a system under study or not depends also on whether the second law of thermodynamics holds or not when all the thermodynamic states are described in these variables only.[76] The *second law of thermodynamics*[r] is formulated as follows:

> (i)[q] For a given macroscopic system, if we choose two equilibrium states specified by energy and work coordinates, there is an adiabatic process at least from one state to the other.
>
> (ii) (Planck's principle) When work coordinates are changed adiabatically, if all of them agree before and after the change, the internal energy of the system cannot decrease.
>
> Here, 'adiabatically' implies that changes are only through mechanical (including electromagnetic) means (i.e., modifying the work coordinates

---

the *thermodynamic space*. Not all the points in this space need be actually realized, but the totality of realizable macrostates make a connected region in this space. It is assumed that the space is equipped with a common sense topology (= the concept of the closeness of points). The thermodynamic space of a composite system is a direct product of thermodynamic spaces for simple systems from which the composite system is made.

[74] ⟪**Thermodynamic processes**⟫ A *process* is a way the states change. Definitions of various thermodynamic processes are different from book to book, so the definitions we adopt are summarized here.

If all the states of the system along a process are infinitesimally close to equilibrium (that is, the process can be described as a path in the system thermodynamic space), it is called a *quasiequilibrium process*. Thermodynamics may be used to calculate the changes of the system along this process. However, the process need not be reversible.

If it is possible to change the final state of a process to its initial state without any trace 'in the world' of the whole round trip, the process is called a *reversible process*. Here, 'any trace 'in the world' ' is a vague statement, but its meaning is that any external systems interacting with the system during the process energetically (and materially) appreciably must return to their states macroscopically indistinguishable from their original states. If any infinitesimal portion of a process is reversible, the process is called a *retraceable process*. This is often informally called a reversible process; this abuse of the terminology seems benign, so this book adopts this abuse.

[75] ⟪**Importance of being quasiequilibrium**⟫ Irrespective of the process being quasiequilibrium or not, if the change is infinitesimal, there must be such a differential form. Why do we have to require the change to be quasiequilibrium? Because, otherwise, $f_i$ cannot be described uniquely in terms of thermodynamic coordinates. If change is rapid, for example, the gas pressure (which may often be mechanically definable even out of equilibrium) cannot be determined by the energy and the volume of the gas.

[76] If we wish to construct thermodynamics axiomatically, variables $E, \{X_i\}$ are primitive concepts without any definition.

alone). It is not the sense appearing in the adiabatic theorem in mechanics.[77]

From these and precise formulation of the zeroth and the fourth laws the existence of entropy $S$ is proved.[78] In particular, it can be proved that $\omega = dE - \sum_i x_i dX_i$ has the absolute temperature as its integrating factor, and $\omega = TdS$.

In a quasistatic reversible process, the *Gibbs relation***r**

$$dE = TdS + \sum_i x_i dX_i \ (= TdS - PdV + xdX + \cdots) \tag{1.3.2}$$

holds. Here, $P$ is pressure, $V$ is volume, and $x, X$ denotes the generic conjugate thermodynamic variable pair (if the amount of matter $N$ may be varied, $\mu dN$ is added, where $\mu$ is chemical potential; this will be discussed in [T10]). From the statistical mechanics point of view, the Gibbs relation is more convenient in the following form regarding entropy as the basic quantity:

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{x}{T}dX + \cdots. \tag{1.3.3}$$

What the fundamental laws assert is that entropy never decreases in an adiabatic process. If the system energy is changed only by modifying the work coordinates, entropy can never be decreased. Therefore, if starting from an equilibrium state, a system reaches another equilibrium state through varying work coordinates only, generally the system entropy increases. This is the principle of entropy increase in thermodynamics. However, this does not mean that entropy increases along the process as a function of time. Thermodynamic entropy is defined only for equilibrium states.

---

[77] **⟪Adiabatic process⟫r** It is not very easy to clearly specify what the adiabatic process is when we do not clearly know what 'heat' is. According to the common sense, 'it is a process realized in a Dewar jar,' but what is required is the absence of heat exchange; a system can be thermodynamically adiabatic, even if it is in contact with a heat bath. Suppose a system can be changed from a state A to state B under isolation solely with the mechanical changes of work coordinates. If we remove the isolation condition and can still bring the state A to B by exactly the same mechanical changes of work coordinates, the process is called an adiabatic process. **r**Using thus defined adiabatic process, we can define heat unambiguously. After defining heat precisely, we can generalize the definition of the adiabatic process**r** as a process that does not have any net heat exchange with the environment.

Incidentally, as to quantum mechanical adiabatic process see T. Kato, "On the adiabatic theorem of quantum mechanics," J. Phys. Soc. Jpn., **5**, 435 (1950). The paper contains an ultimate form of the adiabatic theorem. Quantum-mechanical adiabatic process is a thermodynamic adiabatic process, but the converse is totally false.

[78] **⟪If the reader is serious about thermodynamics⟫** The discussion here is not very rigorous, so those who are interested in precise formulation of thermodynamics, read E. H. Lieb and J. Yngvason, "The physics and mathematics of the second law of thermodynamics," Phys. Rep., **310**, 1-96 (1999). A good introductory summary by the same authors may be found in Notices Am. Math. Soc. 45, 571 (1998).

In the usual textbooks of statistical mechanics, statistical mechanics is completed by imposing a statistical law (the principle of equal probability) on the set $\omega(E, \{X_i\})$ of microstates consistent with a macrostate $(E, \{X_i\})$. Actually, however, if we have a complete set of the correspondence rules that provide the translation of thermodynamic variables in terms of mechanically meaningful quantities, the fundamental purpose of statistical mechanics, i.e., deriving thermodynamics from the microscopic description of the system, is fulfilled. The first part of the correspondence rule is, as stated already, about the quantities whose natural interpretation in terms of mechanics is without controversy. The correspondence rule is not complete with this part only, because thermodynamics discusses thermal quantities as well. Therefore, as the thermal part of the correspondence rule, we adopt *Boltzmann's principle*. If the microscopic dynamics is quantum mechanics, the principle reads:

> Suppose a(n isolated) macrosystem is in a thermodynamic state specified by the thermodynamic coordinates $(E, \{X_i\})$. Let $w(E, \{X_i\})$ be the number of microstates realized in this macrosystem with internal energy $E$ and work coordinates $\{X_i\}$. Then, the entropy of this macrostate is given by
> $$S = k_B \log w(E, \{X_i\}), \tag{1.3.4}$$
> where $k_B$ is Boltzmann's constant: $k_B = 1.3806503 \times 10^{-23}$ J/K.

When Boltzmann proposed this translation rule in 1877, there was no quantum mechanics. If the microscopic dynamics is classical, the above rule reads:

> Suppose a(n isolated) macrosystem is in a thermodynamic state specified by the thermodynamic coordinates $(E, \{X_i\})$. Let $\tilde{w}(E, \{X_i\})$ be the phase volume of the set of microstates realized in this macrosystem with internal energy $E$ and work coordinates $\{X_i\}$.[79] Then the entropy of this macrostate is given by
> $$S = k_B \log \tilde{w}(E, \{X_i\}). \tag{1.3.5}$$

Thus, entropy as a function on the thermodynamic space is given solely in terms of mechanically meaningful quantities (quantities definable within mechanics).

Let us summarize the value of the Boltzmann constant $k_B$.[80]

$$k_B \;=\; 1.3806503 \times 10^{-23} \text{ J/K}$$

---

[79] As we will see later, $\tilde{w}$ is not the phase volume itself, but with a prefactor to make classical mechanics-quantum mechanics correspondence quantitative and with a combinatorial factor to resolve the Gibbs paradox (see Sect. 1.4 (3)). As we will learn soon from the explanation of Boltzmann's original paper, the original proposal did not have such factors.

[80] 10pN is a typical force felt or exerted by molecular machines; a few nm is a typical displacement

$$\begin{aligned} &= 1.3806503 \times 10^{-2}\ \mathrm{pN \cdot nm/K} \\ &= 8.617343 \times 10^{-5}\ \mathrm{eV/K}. \end{aligned}$$

The *gas constant* $R$ is defined by

$$R \equiv N_A k_B = 8.314462\ \mathrm{J/mol \cdot K} = 1.986\ \mathrm{cal/mol \cdot K}. \tag{1.3.6}$$

Here, $N_A = 6.02214078(18) \times 10^{23}/\mathrm{mol}$[81] is *Avogadro's constant* and 1 cal = 4.18605 J. How can we experimentally determine $k_B$? Can we do it with equilibrium observations only?

It is convenient to remember that at room temperature (300 K):[82]

$$\begin{aligned} k_BT &= 4.14\ \mathrm{pN \cdot nm} \\ &= 0.026\ \mathrm{eV}, \\ RT &= 0.6\ \mathrm{kcal/mol}. \end{aligned}$$

If entropy $S$ is determined as $S = f(w)$ solely by the number $w$ of microstates consistent with a macrostate, $S$ must have the form of (1.3.4). Let us divide the macroscopic system into two macroscopic subsystems, 1 and 2. If $S_i$ is the entropy of subsystem $i$, then thermodynamics (the fourth law) tells us

$$S = S_1 + S_2. \tag{1.3.7}$$

On the other hand, these two subsystems may be assumed to be statistically independent without contradicting thermodynamics (thanks to the zeroth law), so if $w_i$ is the number of microstates consistent with the macrostate of subsystem $i$, then $w = w_1 w_2$. (1.3.7) reads

$$S = f(w) = f(w_1 w_2) = S_1 + S_2 = f(w_1) + f(w_2). \tag{1.3.8}$$

---

of molecular motors. Cf., the diameter of DNA is 2 nm (its pitch is 3.4 nm); the $\alpha$-helix pitch is 3.6 amino acid residues = 0.54 nm. To ionize an atom, a few electron volts are needed, so, if $T$ is the room temperature (300 K), it is about 100 $k_BT$. Note that even on the surface of the sun (with the temperature corresponding to the black body radiation of 6000 K), hydrogen atoms are not significantly ionized.

[81] **⟪Definition of the Avogadro constant⟫** This is defined as the number of atoms in a 0.012 kg of $^{12}$C. The Avogadro constant value is due to B. Andreas et al., "Determination of the Avogadro constant by counting the atoms in a $^{28}$Si crystal," Phys. Rev. Lett., **106**, 030801 (2011).

[82] Under physiological condition, hydrolysis of a single ATP molecule provides about $20k_BT$.

Let us assume that $f$ is a monotone function of $w$. Then, $S$ must be proportional to the logarithm of $w$:

$$S = k_B \log w. \tag{1.3.9}$$

Thus, the reader might have thought Boltzmann's principle is almost trivial, but the core of Boltzmann's real insight is that $S$ is determined by the number of the microstates consistent to the macrostate.

**Cauchy's functional equation $f(x+y) = f(x) + f(y)$ ($\forall x, y \in R$)q**
This relation immediately tells us that for any positive integer $n$

$$f(nx) = nf(x). \tag{1.3.10}$$

Therefore, if $m$ is a positive integer,

$$mf(1/m) = f(1). \tag{1.3.11}$$

Combining the above two relations, we obtain

$$f(n/m) = (n/m)f(1). \tag{1.3.12}$$

$f(-n/m) = (n/m)f(-1)$ can also be obtained. Since $f(-1) + f(1) = f(0) = 0$, we conclude that $f(x) = cx$ if $x$ is rational for some constant $c$. To relate the values at rational $x$ and irrational $x$, we need extra conditions. If we require continuity, we can conclude that $f(x) \propto x$ for any real number $x$. The requirement of monotonicity also implies the same conclusion.

Even though we say 'statistical' mechanics, in the above we never encountered probability nor statistics. Then, in conventional introduction of statistical mechanics why does a statistical principle called the "principle of equal probability" come to the foreground instead of the correspondence rules? To seek an answer, let us review how Boltzmann reached his correspondence (1.3.5).

**How did Boltzmann reach his principle? (and its answer)**
Let us taste the original paper.[83] Main steps are stated as questions and answers to them. Let us consider a gas consisting of $N$ particles in a container with volume $V$. Let $w_n$ be the number of particles with the (one-particle) energy between $(n-1)\varepsilon$ and $n\varepsilon$ ($\varepsilon > 0$). Thus, the set $\{w_n\}$ specifies a collection of microstates of the system with $w_n$ particles in the one particle energy bin with the energy in $((n-1)\varepsilon, n\varepsilon]$.
(1) Show that maximizing the number of ways ('Komplexionszahl') to realize a collection of microstates ('Komplexion') specified by $\{w_n\}$ is equivalent to the minimization condition for

$$M = \sum w_n \log w_n. \tag{1.3.13}$$

---

[83] L. Boltzmann, "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechinung respective den Sätzen über Wärmegleichgewicht," Wiener Berichte **76**, 373 (1877) ("On the relation between the second law of thermodynamics and probability calculation concerning theorems of thermal equilibrium") [Accession of Queen Victoria to 'Empress of India' was proclaimed at the Delhi Durbar of 1877.]

The Komplexionszahl reads

$$Z_K = \frac{N!}{w_1!w_2!\cdots w_n!\cdots}, \tag{1.3.14}$$

so maximizing this is equivalent to minimizing the denominator or its logarithm:

$$\log(w_1!w_2!\cdots w_n!\cdots) = \sum_n \log w_n! = \sum (w_n \log w_n - w_n) = M - N. \tag{1.3.15}$$

The original paper kindly discusses that we can discard numerical factors, etc., in Stirling's formula.
(2) Write $w_i = w(x)\epsilon$ and simultaneously take the $n \to \infty$ and $\varepsilon \to 0$ limits, maintaining $x = n\varepsilon$ finite. Show that minimizing $M$ is equivalent to minimizing

$$M' = \int w(x) \log w(x) dx. \tag{1.3.16}$$

Substituting the quantities in $M$ as indicated, we have

$$\sum_n w_n \log w_n = \sum_n w(x)\varepsilon \log[w(x)\varepsilon] = \sum_n w(x)\varepsilon \log w(x) + \sum_n w(x)\varepsilon \log \varepsilon \tag{1.3.17}$$

The first term is a Riemann sum, so we obtain (1.3.16). The second term is $N \log \varepsilon$ and is unrelated to the number of complexions, so we may ignore it.
(3) We should not ignore the constraints that the total number of particles is $N$ and the total energy is $E$. Under this condition, derive Maxwell's distribution in 3-space by minimizing $M'$.

In the original paper Boltzmann regarded the variable $x$ as the three components of velocity vector $v_x, v_y, v_z$. Here, we take the momentum $\boldsymbol{p}$ and the position $\boldsymbol{r}$ as $x$: $w(x) = w(\boldsymbol{r}, \boldsymbol{p})$. The constraints are

$$N = \int d\boldsymbol{r}d\boldsymbol{p}\, w(\boldsymbol{r}, \boldsymbol{p}), \quad E = \int d\boldsymbol{r}d\boldsymbol{p}\, w(\boldsymbol{r}, \boldsymbol{p})E(\boldsymbol{p}), \tag{1.3.18}$$

where $E(\boldsymbol{p}) = \boldsymbol{p}^2/2m$ is the energy of a single particle state $\boldsymbol{p}$ with $m$ being the mass of a gas particle. Using Lagrange's technique, we should maximize ($\alpha$ and $\beta$ are multipliers)

$$\int d\boldsymbol{r}d\boldsymbol{p}\, w(\boldsymbol{r}, \boldsymbol{p})[\log w(\boldsymbol{r}, \boldsymbol{p}) + \alpha + \beta E(\boldsymbol{p})]. \tag{1.3.19}$$

Hence, ($E = (3/2)Nk_BT$ is used to fix $\beta$)

$$w(\boldsymbol{r}, \boldsymbol{p}) = \frac{N}{V} \frac{1}{(2\pi m k_B T)^{3/2}} e^{-p^2/2k_BTm}. \tag{1.3.20}$$

Thus, we have obtained the Maxwell distribution.
(4) Now, Boltzmann realized that $\log Z_K = \log N! - M'$ gives the entropy of the ideal gas. Based on this finding, he proposed

$$S \propto \log(\text{Number of 'Komplexions'}). \tag{1.3.21}$$

Compare this and the formula for $S$ obtained thermodynamically, as Boltzmann did, to confirm his proposal.

If we compute (1.3.21) (i.e., $N \log N - M'$) with the aid of $w$

$$S = N \log V + N \left( \frac{3}{2} + \frac{3}{2} \log(2\pi k_B T m) \right) = N \log V T^{3/2} + \text{ const.} \tag{1.3.22}$$

This agrees with the entropy obtained thermodynamically (apart from $\log N!$; the formula is not extensive). Indeed, the Gibbs relation is

$$dS = \frac{1}{T} dE + \frac{P}{T} dV, \tag{1.3.23}$$

so with the aid of the internal energy $E = (3/2)Nk_B T$ and the equation of state, we have

$$S = Nk_B \log T^{3/2} + Nk_B \log(V/N) + \text{ const.} \tag{1.3.24}$$

Usually, the story ends here (so did Boltzmann's original paper). However, being a much deeper thinker than is usually regarded, for a macrosystem described by $E$ and $V$ Boltzmann confirmed that his formula of entropy (1.3.5) satisfied the Gibbs relation for general classical many-body systems; in particular, $(dE + PdV)/T$ is a complete differential. Boltzmann called a distribution on the totality of microstates consistent with a macrostate an *orthodic distribution*, if it gives a statistical theory consistent with thermodynamics. Thus, Boltzmann proved that microcanonical ensemble is an orthodic ensemble. A problem at the end of this chapter will discuss the demonstration of the Gibbs relation for the microcanonical ensemble. The most fundamental problem of statistical mechanics may be the characterization of orthodic distributions: what characterizes the distributions of microstates consistent with thermodynamics?

Does a statistical assumption or a statistical principle appear in Boltzmann's logic? It is latently in the starting point of the argument (1) in the above fine-lettered entry. Its central idea is that a macrostate consistent with maximal number of microstates is observed as a thermodynamic equilibrium state. It says that the number is important, so it is a declaration of the idea that all the microstates are counted equally. Namely, the principle that all the microstates consistent with a macrostate have the same weight (the principle of equal weight or probability) is assumed implicitly (probably unconsciously). This suggests that if we combine thermodynamics and a statistical principle, we could derive Boltzmann's principle, and formulate statistical mechanics. That is, the correspondence rule for thermodynamic coordinates + a certain statistical principle (that imposes a distribution on the set $\omega(E, \{X_i\})$) should be able to formulate statistical mechanics.

In this way, we can understand the (at least potential) importance of the statistical principle. There are many authors who assert that such a key principle must be derivable from the mechanics of the microscopic system (if thermodynamic coordinates are translated appropriately in terms of mechanics). Assuming an appropriate translation rule, let us explore whether such an assertion is sensible or not.

In classical mechanics, the distribution is a probability measure[84] on the phase space spanned by all the position and momentum coordinates of the particles. Since each point in the phase space corresponds to a single microstate, the distribution assigns statistical weights to individual microstates, and we can imagine the distribution as a cloud in the phase space. The equilibrium distribution must be time-independent (must correspond to an invariant measure).[85] In quantum mechanics the distribution (or the *density operator*) is a linear operator $\rho$ satisfying the following two conditions:
(i) Normalization condition: $Tr\rho = 1$,
(ii) If $A$ is non-negative definite,[86] $Tr\rho A \geq 0$.[87]
For example,**r**

$$\rho = \sum_k P_k |k\rangle\langle k|, \tag{1.3.25}$$

where the summation is over an appropriate set of orthonormal kets in the vector space needed to describe the system and $\sum P_k = 1$ ($P_k > 0$), is a respectable density operator. If these kets are eigenkets of the system Hamiltonian (eigenstates of the system Hamiltonian), it is invariant in time. How should we choose $P_k$'s? We know nothing more than their time-independence and non-negativity. We encounter similar non-uniqueness in classical mechanics. That is, the time-invariance condition falls short of specifying the distribution we need.

What determines an invariant distribution? If we start from an initial distribution and run the system long enough, usually it settles down to an invariant measure.[88] Therefore, selecting a 'right' distribution is equivalent to selecting a 'right' initial condition. However, everybody knows that an initial condition is chosen by the user of the mechanics, not by the equation of motion itself. Nothing in classical or

---

[84] This may simply be understood as a weight whose total mass is normalized to 1.

[85] To be precise, the distribution is required to give time-independent results only for thermodynamic observables, so the distribution itself need not be invariant. Thus, for example, we have only to demand 'weak' time invariance for the expectation values of the jumpsum1sum functions (roughly, it is a function of many variables expressed as a sum of many functions with a few variables) introduced by Khinchin. No serious attempts in this direction seem to exist, so this book simply demands invariance in the usual sense.

[86] This means $\langle\ |A|\ \rangle \geq 0$ for any ket $|\ \rangle$.

[87] A linear operator satisfying (ii) is self-adjoint.

[88] If a system is chaotic enough, loosely put. This is the classical case. For quantum systems, the dynamics is linear, so the initial condition is everything.

quantum mechanics tells us how to choose initial conditions, so we have to conclude that the distribution needed to develop statistical mechanics cannot in principle be deduced from mechanics alone.[89] In Section 1.1, we thought what statistical principle was convenient. Here, since mechanics cannot derive any principle, we follow the standard approach of assuming a convenient principle.

As the *statistical principle* we adopt the *principle of equal probability*:

> All the microstates consistent with the thermodynamic coordinates specifying a macrostate have the same statistical weight.

In other words, all the microstates in $\omega(E, \{X_i\})$ are sampled with an equal probability.[90] The distribution of microstates specified by this principle is called the *microcanonical distribution*.

Quantum mechanically, the sum defining the density matrix (1.3.25) is over all the microstates consistent with the macrostate[91] with $P_k$ being equal for all such microstates. The microcanonical density operator is written as the following normalized sum

$$\rho = \frac{1}{w(E, \{X_i\})} \sum_k |k\rangle\langle k|, \tag{1.3.26}$$

where the sum is over all the eigenstates with energy eigenvalues between $E - \delta E$ and $E$, and (the expectation values of) work coordinates between $X_i - \delta X_i$ and $X_i$.[92]

---

[89] A common objection to this conclusion is that if almost all initial conditions go to a particular invariant measure, then the invariant measure describes the equilibrium distribution, so we are allowed to claim that mechanics dictates the statistical measure. The error in this argument is in the uncritical expression "almost all initial conditions." With respect to what sampling method (measure) is this statement made? We must choose the observation measure of the initial condition. This choice is beyond mechanics, needless to say.

[90] ⟪**Significance of the principle of equal probability**⟫ Those who have read Sect. 1.1 may interpret the statistical principle as a means to specify a filter to obtain thermodynamic observables. This is not a postulate that actually such a distribution really exists in the equilibrium state specified by the thermodynamic coordinates. The equal probability may or nay not be the actual distribution.

[91] ⟪**The significance of thermodynamic coordinates**⟫ This was already emphasized, but here another importance must be pointed out. We must not forget that macrostates are uniquely specified only by extensive variables (densities). This must be clear if we imagine a system with phase coexistence (see Chapter 5). For example, if a macroscopic state of an isolated magnet is specified by a temperature below its critical temperature (or has an internal energy corresponding to such a temperature), its magnetization is not unique. Therefore, if we evenly average all the microstates under this specification of the macrostate, the magnetization is zero. That is, the result does not agree with the actually realizable thermodynamic state. (As we will see in Chapter 5 for a 2D Ising model it is even a state that is never realized macroscopically.)

[92] To be precise, $(E - \delta E, E]$ and $(X_i - \delta X_i, X_i]$ determines the shell.

$\delta E$ and $\delta X_i$ denote small widths from the macroscopic point of view. There is a large latitude in the choice of these widths; if they are not extremely small, any choice will do as is discussed in Section 1.5. The normalization constant $w(E, \{X_i\})$ is the number of all the microstates in this shell. Since neither $E$ nor $X_i$ depends on time, it is obvious that the density operator is time-independent.

Classically, the principle of equal probability implies that we compute the expectation value according to the uniform distribution whose support is the set of microstates consistent with the specified thermodynamic coordinates (with some latitude). The resultant distribution is called the microcanonical distribution. Since the subset of the phase space consistent with a macrostate is invariant, and since the phase volume is invariant (as seen just below), it is an invariant distribution.

**Phase flow is incompressible** **r**(**r**to Jarzynski equality, **r**to virial theorem)
Let us show that the phase volume is invariant classically. This is proved 'officially' by *Liouville's theorem***q** (stating that any canonical transformation preserves the phase volume), but here let us show the incompressibility of the phase flow. If there are $N$ point masses in a three dimensional space, the coordinates of the phase space are $q_{1x}, q_{1y}, q_{1z}, \cdots, q_{Nx}, q_{Ny}, q_{Nz}, p_{1x}, p_{1y}, p_{1z}, \cdots, p_{Nx}, p_{Ny}, p_{Nz}$ ($q_{3y}$ is the $y$-coordinate of the third particle position vector, $p_{5x}$ is the $x$-component of the momentum vector of the 5th particle, and so on). The suffixes denote individual degrees of freedom. The phase flow vector $\boldsymbol{v}$ is a $6N$-dimensional vector $\boldsymbol{v} = (\dot{q}_{1x}, \cdots \dot{q}_{Nz}, \dot{p}_{1x}, \cdots, \dot{p}_{Nz})$ (dots denote time differentiation) consisting of the following components:

$$\frac{dq_s}{dt} = \frac{\partial H}{\partial p_s}, \ \frac{dp_s}{dt} = -\frac{\partial H}{\partial q_s}. \tag{1.3.27}$$

That is,

$$\begin{aligned}\boldsymbol{v} \ = \ & \left(\frac{\partial H}{\partial p_{1x}}, \frac{\partial H}{\partial p_{1y}}, \frac{\partial H}{\partial p_{1z}}, \cdots, \frac{\partial H}{\partial p_{Nx}}, \frac{\partial H}{\partial p_{Ny}}, \frac{\partial H}{\partial p_{Nz}}, -\frac{\partial H}{\partial q_{1x}}, -\frac{\partial H}{\partial q_{1y}}, -\frac{\partial H}{\partial p_{1z}}, \right. \\ & \left. \cdots, -\frac{\partial H}{\partial q_{Nx}}, -\frac{\partial H}{\partial p_{Ny}}, -\frac{\partial H}{\partial p_{Nz}}\right). \end{aligned} \tag{1.3.28}$$

The divergence of the flow vector in $6N$-dimensional space reads

$$div\, \boldsymbol{v} = \sum_s \left(\frac{\partial \dot{q}_s}{\partial q_s} + \frac{\partial \dot{p}_s}{\partial p_s}\right) = \sum_s \left(\frac{\partial}{\partial q_s}\frac{\partial H}{\partial p_s} - \frac{\partial}{\partial p_s}\frac{\partial H}{\partial q_s}\right) = 0. \tag{1.3.29}$$

Notice that even if $H$ is explicitly time-dependent, the incompressibility of the phase flow is intact.

If we use the correspondence (translation) rule for thermodynamic coordinates and the principle of equal probability, we may be able to derive Boltzmann's principle from the relation between entropy and heat $dS = d'Q/T$. This is essentially a rehash of Einstein's approach to derive the canonical distribution from thermodynamics.

The microcanonical distribution is[93]

$$w = \int d\Gamma \, \delta(H - E), \tag{1.3.30}$$

This is the classical case but the following argument goes through for quantum mechanics as well by replacing the phase integral with trace. Let us assume that the Hamiltonian contains a parameter $\lambda$ that can be controlled externally. The change of the Hamiltonian due to the change of the parameter $\lambda \to \lambda + d\lambda$ is identified with work by Einstein. Taking account of the internal energy $E$ being also changed, we obtain

$$0 = d \int d\Gamma \frac{1}{w} \delta(H - E) = -d[\log w] + \frac{\int d\Gamma \, (d'W - dE)\delta'(H - E)}{\int d\Gamma \delta(H - E)}. \tag{1.3.31}$$

Since almost all the microstates gives the same thermodynamical results, $d'W - dE$ is almost always constant. Therefore, we obtain

$$d \log w = (d'W - dE) \frac{\int d\Gamma \, \delta'(H - E)}{\int d\Gamma \, \delta(H - E)}. \tag{1.3.32}$$

Since $d'W - dE = -d'Q$, and

$$\left. \frac{\partial \log w}{\partial E} \right|_X = - \frac{\int d\Gamma \, \delta'(H - E)}{\int d\Gamma \, \delta(H - E)}, \tag{1.3.33}$$[94]

we get

$$d \log w = \beta d'Q. \tag{1.3.34}$$

Comparing this with thermodynamic relations, Boltzmann's relation (**??**) has been derived.

In summary,
(i) A method to construct statistical mechanics is to impose the correspondence rules for thermodynamic coordinates and entropy. No statistical principle is required.
(ii) Another approach is to use the correspondence rule for thermodynamic coordinates, but the principle of equal probability is used instead of the correspondence rule for entropy.

---

[93] Here, the traditional formula is written, but precisely speaking, not only the energy (internal energy) but also work coordinate values must be specified, which are not written explicitly.

[94] For the classical ideal gas, we can explicitly compute (1.3.33) and confirm its identity with $\beta$.

The construction of statistical mechanics explained above applies both to isolated and thermally isolated systems.**r** As noted already, isolated systems (for which no exchange of energy is allowed with the external world) are not enough to cover all possible thermodynamic possibilities. Still, usually no separate discussion is found about constructing statistical mechanics for thermally isolated systems. To begin with we must modify the zeroth law ([T6]). Then, statistical mechanics is formulated for thermally isolated systems.

> **[T6] Generalized zeroth law of thermodynamics**
> As noted before, a large subset of the thermodynamic states cannot be realized without interactions with the external environment, so the zeroth law stated only for isolated systems [T1] is not sufficient even for thermodynamics itself. What sort of interactions with environment should be allowed? A macroscopic time-independent setup that maintains (the expectation value[95] of) a particular work coordinate (or more precisely the expectation value of its density) at a chosen value must be tolerated.
>
> Let us introduce the concept of 'thermally isolated system.' It is a system first isolated in the ordinary sense, and then appropriate time-independent mechanical interactions are imposed to keep work coordinates at their specified values. The zeroth law applicable to thermally isolated systems may be formulated as:
>
> (1) If a thermally isolated system is kept in a constant environment for a sufficiently long time, it reaches a state without any macroscopic time dependence. This is called an equilibrium state.
>
> (2) If the obtained equilibrium state is divided roughly evenly with a plane,
>   (i) Each half can be kept in the same equilibrium state in thermal isolation,
>   (ii) if these halves are joined, then the resultant thermodynamic state is an equilibrium state indistinguishable from the original thermodynamic state in thermal isolation.

To construct statistical mechanics, we need mechanical translation of work coordinates. No modification is needed for this. We generalized the Boltzmann principle, that is the mechanical translation of entropy (1.3.4) or (1.3.5). For simplicity, let us assume that the thermally isolated system we consider is an isolated system with the imposed conjugate field $h$ to $X$. The natural thermodynamic potential is the generalized enthalpy $J = E - hX$;**r** $(J, h, \{Y\})$ is the natural thermodynamic variables for the system, where $\{Y\}$ is the work coordinates other than $X$. The system is described as a mechanical system with the Hamiltonian $H = H_0 - h\hat{X}$, where $H_0$ is the system Hamiltonian in isolation and $\hat{X}$ is the mechanical observable corresponding to the work coordinate $X$, with the auxiliary conditions (boundary conditions) corresponding to the experimental condition to prepare this equilibrium state. Then the entropy of this state is given by

$$S(J, h, \{Y\}) = k_B \log w(J, h, \{Y\}). \tag{1.3.35}$$

---

[95] In the thermodynamic limit, the actual value and the expectation value are identical.

where $w(J, h, \{Y\})$ is the number of eigenstates of $H$ belonging to $(J - \delta J, J]$ with a reasonable width $\delta J$ with the specification of the expectation values of observables $\{\hat{Y}\}$ corresponding to the work coordinates $\{Y\}$. [That is, (1.3.35) is exactly the same as the Boltzmann's formula except for replacing $H_0$ with $H$.]

We know the Gibbs relation now reads

$$dJ = TdS - Xdh + \cdots, \tag{1.3.36}$$

so

$$dS(J, h, \{Y\}) = \frac{1}{T}dJ + \frac{X}{T}dh + \cdots. \tag{1.3.37}$$

Notice that this is not a Legendre transformed version of $S$,[96] but since we can obtain $E = J - hX$ and $X$ is specified, $S$ may be understood as a function of the thermodynamic coordinates $(E, X, \{Y\})$, and from (1.3.37) we get consistently

$$dS(E, X, \{Y\}) = \frac{1}{T}dE - \frac{h}{T}dX + \cdots. \tag{1.3.38}$$

## 1.4 Implication of the principles

To construct statistical mechanics we have reached the conclusion that we may adopt (A) or (B) (as postulates):

For a thermally isolated system, besides the correspondence rules for work coordinates:

(A) Entropy is given by $S = k_B \log w$, where $w$ is the number of microstates allowed to the system.

Here, the natural thermodynamic variables are the generalized enthalpy $J = E - \sum_i h_i X_i$, imposed field $h_i$ conjugate to $X_i$ and the rest of the work coordinates $Y_j$. The other choice is

(B) [Principle of equal probability] The probability to observe an event is proportional to the number of microstates compatible with the event.

(B) implies (A), and to do thermodynamics, (A) is enough.

General consequences of the fundamental principles include:

(1) A thermodynamic equilibrium state is a macrostate that maximizes the number of consistent microstates.

---

[96] Notice that the second term is not $Xd(h/T)$.

(2) The number of microstates consistent with an equilibrium state is $O[e^{cN}]$.
(3) Indistinguishability of identical particles.
(4) The equilibrium condition between two macroscopic objects that are allowed to exchange extensive quantities is the agreements of the corresponding conjugate intensive variables.
(5) Absence of macroscopic motions other than overall translation and rotation of macroscopic objects in equilibrium.
(6) Positivity of absolute temperature.

(1) "Maximum entropy principle"
Thermodynamic entropy is defined only for equilibrium states.[97] Therefore, it is a nonsensical statement that an isolated macroscopic system specified by $(E, \{X_i\})$ takes the maximum entropy in equilibrium; thermodynamic entropy is a function of $(E, \{X_j\})$, so its value is uniquely fixed. Thermodynamic entropy is defined only for equilibrium systems in any case.

Let us consider a composite system consisting of two simple systems. Although the subsystems must be in equilibrium individually, they need not be in equilibrium mutually because of the constraints defining the composite system (say, an adiabatic wall between them). Therefore, even though the system is not in equilibrium as a whole without constraints, still we can obtain the entropy of the whole system. The number $w$ of microstates of the composite system may be written as $w = w_1 w_2$, where $w_1$ and $w_2$ are numbers of states of individual simple systems. If some constraints are removed, then more microstates should be allowed to the composite system. That is $w$ must increase. Consequently, the newly reached equilibrium state must have larger entropy than before removing the constraint. This is the thermodynamic entropy maximum principle. This implies, as shown below with fine letters, that entropy is a convex upward function of the thermodynamic coordinates.

**[T7] Entropy maximum principle in thermodynamics**
This principle is about the final equilibrium state that can be reached from an initial equilibrium state of an isolated system when some constraints in the initial state are removed. A precise statement is as follows: let a system be a composite system consisting of simple systems $i$ $(i = 1, \cdots, N)$. If each simple system is in equilibrium the entropy of the composite system is a sum of the entropies of simple systems $S^{(i)} = S^{(i)}(E^{(i)}, \{X_j^{(i)}\})$, where superscripts $(i)$ indicate simple systems. When a composite system is formed from these simple systems, we may allow some exchange of work coordinates and energy among them, so the initial isolation condition for each simple system may be relaxed. Eventually, the composite system reaches an equilibrium state. Each subsystem is also in equilibrium so its entropy may be expressed as $S^{(i)}$.

---

[97] Practically, if a macrostate is close to equilibrium, there are several natural ways to extend the definition of entropy. However, at present they are pragmatic and not of fundamental nature.

The compound system with a relaxed subsystem relations is in equilibrium, when

$$\hat{S} = \sum_i S^{(i)}(E^{(i)}, \{X_j^{(i)}\}) \tag{1.4.1}$$

is maximized under the relaxed condition.

The concavity (convexity upward) of entropy follows from this. Suppose two (simple) systems made of the same material in equilibrium states $\omega^{(1)} = (E^{(1)}, \{X_i^{(1)}\})$ and $\omega^{(2)} = (E^{(2)}, \{X_i^{(2)}\})$, respectively, are in contact. Throughout the process we assume that the whole system is surrounded by adiabatic walls. Then, eventually a single equilibrium state $\omega = (E, \{X_i\}) \equiv (E^{(1)} + E^{(2)}, \{X_i^{(1)} + X_i^{(2)}\})$ would be formed (Fig. 1.4.1).

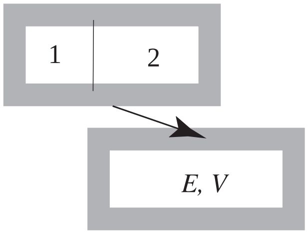


Fig. 1.4.1 The surrounding wall is adiabatic. The two equilibrium states 1 and 2 are allowed to make any interactions in the container. Eventually, the composite system will reach an equilibrium state specified by the total amounts of extensive quantities.

Let us be more explicit with the independent extensive variables to be internal energy and volume. The entropy maximum principle tells us that

$$S(E, V) = \sup[S(E_1, V_1) + S(E_2, V_2)], \tag{1.4.2}$$

where the supremum is taken under the condition $E = E_1 + E_2$ and $V = V_1 + V_2$. Here, since both systems are assumed to be made of the same substance, the use of the same symbol $S$ for systems 1 and 2 means that the functional form of $S$ is identical. This implies with the aid of the extensivity of $S$**r**

$$\begin{aligned} S((1-\alpha)E + \alpha E', (1-\alpha)V + \alpha V'') \\ \geq (1-\alpha)S(E, V) + \alpha S(E', V') \end{aligned} \tag{1.4.3}$$

for any $\alpha \in [0, 1]$, that is, $S$ is a concave function (its graph is convex upward). This inequality is, as we will learn later, a manifestation of the stability of the equilibrium state.

### (2) Impossibility (unrealistic nature) of mechanical adiabatic process

Since entropy is extensive, if $N$ is the number of particles in the system, $\log w = S/k_B \propto N$ and the dimensionless proportionality constant $c$ for this relation is of order unity. That is,

$$w \sim e^{cN + o[N]}. \tag{1.4.4}$$

Generally speaking, the spacing of energy levels of a macroscopic system is of order

$10^{-cN}$ (here the unit of energy may be anything macroscopic, say, J or erg).[98] The energy levels of a macrosystem is so densely packed that it is impossible to realize a process that may satisfy conditions for the adiabatic theorem of mechanics.[99]

(3)**r** Indistinguishability of identical particles (Gibbs "paradox")
Suppose the statistical weight $W$ of a system of volume $V$ consisting of $N$ almost independent particles is given by

$$W(V, N) = V^N f(N). \tag{1.4.5}$$

The Boltzmann principle and the extensivity of entropy imply $W(rV, rN) = [W(V, N)]^r$ for any positive real $r$.[100] That is (note that $V$ is extensive),

$$(rV)^{rN} f(rN) = V^{rN} f(N)^r, \tag{1.4.6}$$

or

$$f(N)^r = f(rN) r^{rN}. \tag{1.4.7}$$

As a mathematical problem, $r$ can be anything, so setting $r = 1/N$, we get $f(N)^{1/N} = f(1)/N$. That is, with the aid of Stirling's formula $N! \sim (N/e)^N$, we conclude**r**

$$f(N) = a^N/N! \tag{1.4.8}$$

---

[98] Is this correct for ideal gas? The statement is true if the system is without extremely high symmetry. Let $\{\varepsilon_i\}$ be the totality of the one-particle energy eigenvalues. If two such particles interact with each other, each level splits into two. If there are $N$ particles, each level would split into $2^N$ levels. That is, if there is no degeneracy, the number of levels increase exponentially with the number of particles $N$. Consequently, the energy spacings decrease exponentially as the number of degrees of freedom increases.

[99] More precisely, we should state as follows: mechanical adiabatic process is clearly a thermodynamic quasistatic process. Therefore, if entropy were to increase under mechanical adiabatic condition, it implies that mechanics contradicts thermodynamics, and the idea of statistical mechanics is meaningless. On the other hand, since mechanical adiabatic process is an extremely special quasistatic process, the invariance of entropy under mechanical adiabatic process means almost nothing to thermodynamics. We should clearly recognize that macroscopic quasiequilibrium processes are much faster and violent processes than the ones that allow the adiabatic theorem (in mechanics) to hold.

[100] **⟪Irrelevance of particle indistinguishability in computer experiments⟫** As the reader can see from the argument around here, if the number of the particles is kept constant, there is no way to determine $f$ (furthermore, it is a constant, so we may simply ignore it). That is, there is no thermodynamic observable that are affected by the (in)distinguishability of particles, if there is no change of the number of particles in the system. Therefore, for example, in the computer experiment without particle number change we have nothing to worry about particle (in)distinguishability. See the last part of Sect. 2.1

for large $N$, where $a$ is a positive constant. Therefore, the statistical weight must have the following general form

$$W = \frac{1}{N!} z^N. \tag{1.4.9}$$

Here, $z$ may be interpreted as the statistical weight for each (weakly interacting) microscopic entity.

The factor $N!_{\mathbf{r}}$ indicates that the classical atomism (naive realism) is not consistent with thermodynamics; we must regard individual but identical entities indistinguishable when we compute the combinatorial factor. This is called *Gibbs' paradox.*[101] Intuitively, this may correspond to interpreting particles as 'pixels' on a computer screen that are turned on.

It is important to recognize that Gibbs himself did not regard this as a paradox, because he knew that nothing was known about the physics of microscopic world. A paradox about the reality is always due to our misconception (or wrong metaphysical premise) about the reality.

**Stirling's formula**$_{\mathbf{r}}$
$\Gamma(n+1) = n! \simeq \sqrt{2\pi n}(n/e)^n$ is a very famous formula and is used often in a simplified form $n! \sim (n/e)^n$. Since it is an important formula, a demonstration is given here. Here, we use the method Laplace used to derive the formula. It is a general method to obtain an asymptotic formula called *Laplace's method.* Consider

$$F(\theta) = \int_{-\infty}^{+\infty} e^{\theta h(x)} dx, \tag{1.4.10}$$

where $h$ is a real $C^2$-function with the following properties:
(i) $h(0) = 0$ is an absolute maximum of $h$, and $h < 0$ for any nonzero $x$.
(ii) There are positive constant $a$ and $b$ such that $h \le -a$ for $|x| \ge b$.
We must of course assume that the integral converges for sufficiently large $\theta$. Then, in the $\theta \to \infty$ limit, we get

$$F(\theta) \sim \sqrt{2\pi}(-\theta h''(0))^{-1/2}. \tag{1.4.11}$$

This is a logic used repeatedly in statistical mechanics and sometimes called informally the saddle point method. Formally, this can be obtained by Taylor-expanding $h$

---

[101] ⟪**Gibbs' paradox for mixing**⟫ If we mix two distinguishable ideal gases at the same temperature and pressure, the entropy increases due to mixing beyond the sum of the original entropies of two gases, but if we mix two indistinguishable ideal gases at the same temperature and pressure, the entropy is the same as the sum of the original entropies of two gases. This is also (and probably more often than the ontological paradox in the text) called Gibbs' paradox. See the problem set at the end of the chapter.

around 0 to the second order and by using Gaussian integral.

Now, let us apply this to the Gamma function

$$\Gamma(z+1) = \int_0^\infty dt\, t^z e^{-t}. \tag{1.4.12}$$

Substituting $t = z(1+x)$, we get

$$\Gamma(z+1) = e^{-z} z^{z+1} \int_{-1}^\infty \left[e^{-x}(1+x)\right]^z dx. \tag{1.4.13}$$

$h$ in (1.4.10) reads $-x + \ln(1+x)$ and $\theta = z$, so it satisfies the condition of Laplace's method, and $h''(0) = -1$. Hence, we get

$$n! \sim \sqrt{2\pi} e^{-n} n^{n+1/2}. \tag{1.4.14}$$

The above derivation looks natural. Another interesting approach is to use the MacLaurin summation formula.[102]

(4) Equilibrium condition between two systems in contact**q**

Suppose two macrosystems 1 and 2 are in contact and are isolated as a whole. The total entropy is $S = S_1 + S_2$, so $\delta S = 0$ is the equilibrium condition for the isolated system (it is better to write as $\delta S \le 0$ for any possible variation of variables, because $S$ is maximum). Here, $\delta$ implies virtual variation.[103] To compute the derivatives of entropy, the Gibbs relation for entropy is convenient

$$dS = \frac{1}{T} dE + \frac{P}{T} dV - \frac{\mu}{T} dN - \frac{x}{T} dX, \tag{1.4.15}$$

where $X$ is a generic extensive variable and $x$ is its conjugate intensive variable with respect to energy. Therefore,

$$\frac{1}{T_1}\delta E_1 + \frac{1}{T_2}\delta E_2 + \frac{P_1}{T_1}\delta V_1 + \frac{P_2}{T_2}\delta V_2 - \frac{\mu_1}{T_1}\delta N_1 - \frac{\mu_2}{T_2}\delta N_2 - \frac{x_1}{T_1}\delta X_1 - \frac{x_2}{T_2}\delta X_2 = 0. \tag{1.4.16}$$

Since the system is isolated, $\delta E = \delta E_1 + \delta E_2 = 0$, $\delta V_1 + \delta V_2 = 0$, etc.,

$$\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\delta E_1 + \left(\frac{P_1}{T_1} - \frac{P_2}{T_2}\right)\delta V_1 - \left(\frac{\mu_1}{T_1} - \frac{\mu_2}{T_2}\right)\delta N_1 - \left(\frac{x_1}{T_1} - \frac{x_2}{T_2}\right)\delta X_1 = 0. \tag{1.4.17}$$

$E_1$, $V_1$, etc., can be changed independently, so we obtain $T_1 = T_2$, $P_1 = P_2$, etc., are the equilibrium conditions.

---

[102] See J. Havil, *Gamma, exploring Euler's constant* (Princeton, 2003), which is very enjoyable.

[103] Although it is called 'virtual,' it is actually real; such variations are incessantly realized by thermal noise in the actual system.

What happens if the wall is deformable but not diathermal? If no particle exchange is possible, this gives us $P_1 = P_2$, but the temperatures need not be identical.[104]

(5) The impossibility of macroscopic relative motion in equilibrium[r]
Entropy $S$ is a function of the (expectation value of) internal energy $E$ seen from the center-of-mass co-rotating coordinates of the system (in thermodynamics it is called the internal energy):

$$E = U - \frac{\boldsymbol{P}^2}{2M} - \frac{1}{2}\boldsymbol{L}I^{-1}\boldsymbol{L}, \tag{1.4.18}$$

where $U$ is the total energy, $\boldsymbol{P}$ the total momentum, $\boldsymbol{L}$ the total angular momentum, $M$ the total mass of the system, and $I$ the inertial moment tensor around the center of mass. Let us divide the system into several subsystems. We put suffix $a$ to macroobservables of the subsystem $a$. We have $\boldsymbol{P} = \sum \boldsymbol{P}_a$ and $\boldsymbol{L} = \sum_a \boldsymbol{L}_a$. The angular momentum $\boldsymbol{L}_a$ of the subsystem $a$ around the center of mass of the whole system may be written as $\boldsymbol{L}_a = \boldsymbol{\ell}_a + \boldsymbol{r}_a \times \boldsymbol{P}_a$, where $\boldsymbol{r}_a$ is the relative position vector of the center of mass of subsystem $a$ with respect to the center of mass of the whole system, and $\boldsymbol{\ell}_a$ is the angular momentum of subsystem $a$ around its own center of mass. The Gibbs relation for subsystem $a$ reads (for simplicity, other extensive variables than entropy are suppressed in $dE = TdS + \cdots$; we write $U$ as a function of $S$, momentum and angular momentum)

$$TdS_a + \boldsymbol{v}_a \cdot d\boldsymbol{P}_a + \boldsymbol{\omega}_a \cdot d\boldsymbol{\ell}_a - dU_a = 0, \tag{1.4.19}$$

where $\boldsymbol{v}_a$ is the center of mass velocity of subsystem $a$, and $\boldsymbol{\omega}_a$ is its angular velocity around its center of mass.

The additivity of entropy implies (we use the fact that the total energy is preserved)

$$TdS + \sum_a \boldsymbol{v}_a \cdot d\boldsymbol{P}_a + \sum_a \boldsymbol{\omega}_a \cdot d\boldsymbol{\ell}_a = 0. \tag{1.4.20}$$

---

[104] ⟪**A piston without heat exchange**⟫ In this case energy and volume cannot be changed independently, so the Gibbs relation for entropy is not useful. Rather, $dE = TdS - PdV + \cdots$ is more convenient. $dS = dS_1 + dS_2$, but due to adiabaticity $dS_1 = dS_2 = 0$, so we cannot say anything about temperature. From the volume change, we can conclude $P_1 = P_2$ only. Then, what actually happens, if such a system is left alone for a long time? As the zeroth law asserts, the system eventually reaches a certain equilibrium state, but it cannot be obtained by a simple equilibrium thermodynamic argument alone. We must pay due attention to the nonequilibrium details of the actual process that happens. See the following excellent paper: A. Fruleux, R. Kawai, and K. Sekimoto, "Momentum Transfer in Nonequilibrium Steady States," Phys. Rev. Letters **108**, 160601 (2012).

This can be rewritten as[105]

$$TdS + \sum_a (\boldsymbol{v}_a - \boldsymbol{\omega}_a \times \boldsymbol{r}_a) \cdot d\boldsymbol{P}_a + \sum_a \boldsymbol{\omega}_a \cdot d\boldsymbol{L}_a = 0. \tag{1.4.21}$$

Now, let us write down the condition to maximize the total entropy of the system under the constant total momentum, and total angular momenta. Using Lagrange's multipliers $\boldsymbol{v}$ and $\boldsymbol{\omega}$, we obtain

$$\frac{\partial S}{\partial \boldsymbol{P}_a} = \frac{\boldsymbol{v} + \boldsymbol{\omega}_a \times \boldsymbol{r}_a - \boldsymbol{v}_a}{T}, \tag{1.4.22}$$

$$\frac{\partial S}{\partial \boldsymbol{L}_a} = \frac{\boldsymbol{\omega} - \boldsymbol{\omega}_a}{T}. \tag{1.4.23}$$

Consequently,

$$\boldsymbol{v}_a = \boldsymbol{v} + \boldsymbol{\omega}_a \times \boldsymbol{r}_a, \quad \boldsymbol{\omega}_a = \boldsymbol{\omega}. \tag{1.4.24}$$

That is, in equilibrium, a macroscopic system can move only as a rigid body: no macroscopic relative motion of the parts is allowed.

In the above consideration we have assumed that the momenta are exchanged between the subsystems. Suppose a macroscopic system may be divided into two subsystems that can move relatively without friction. Then, note that the above argument does not work (i.e., without dissipation relative macroscopic motion is allowed even in equilibrium).[106]

(6) The absolute temperature must be positive (for a stable equilibrium state).
Almost as a byproduct of (5) we can conclude that the absolute temperature is nonnegative. Suppose $T$ is negative. Then, $S$ is a decreasing function of the internal energy. Therefore, if $T < 0$, then the entropy can be increased by pumping the energy into the macroscopic degrees of freedom (see (1.4.18)).[107] Consequently, the system explodes. Therefore, $T$ cannot be negative.

However, if the system internal energy is bounded from above (needless to say, no translational nor rotational motion should be allowed), then it cannot increase indefinitely and the above argument does not apply (as we will see later for the spin system not interacting with the lattice degrees of freedom).

---

[105] We have used the following relation: $\boldsymbol{A} \cdot \boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{B} \cdot \boldsymbol{C} \times \boldsymbol{A} = \boldsymbol{C} \cdot \boldsymbol{A} \times \boldsymbol{B}$, where vectors are 3-vectors and these products describe the volume of the parallelepiped spanned by $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ (in this order; if we exchange two variables, the volume is 'inverted' and the sign flips).

[106] However, no local velocity fluctuation can be generated thermally without dissipation. Nothing without any decaying process can be produced spontaneously in the stable world.

[107] Small scale organization of microscopic thermal noise into mesoscopic scale collective motion always occurs as fluctuation, but this does not cause anything serious, because the system is stable.

We are tempted to conclude that in the $T \to 0$ limit $S = 0$, because any ground state of an ordinary quantum mechanical system is believed to be non-degenerate. The third law of thermodynamics (see [T8]) is a thermodynamic claim, so it is for a very large system. That is, the thermodynamic limit must be taken *before* $T \to 0$ limit. Then, we easily realize that the third law is not the same as the non-degeneracy of the ground state. Just as we have experienced with the adiabatic theorem, it is dangerous to translate mechanics propositions directly into thermodynamic propositions.

**[T8] The third law of thermodynamics**$_{\mathbf{r}}$

In 1902 Richards[108] investigated electrochemical potentials at low temperatures, and found that the change in the Gibbs free energy $\Delta G$ converges to that of enthalpy $\Delta H$ in the $T \to 0$ limit. Based on his study and other low temperature chemical reaction data, in 1906,[109] Nernst[110] proposed that $\Delta S = 0$ for any reversible change in the $T \to 0$ limit. In other words, he proposed that the entropy of a substance in equilibrium is bounded and continuous at $T = 0$. A stronger form of Nernst's law is that the derivative of $S$ vanishes at $T = 0$. Later, Planck proposed that the entropy of a pure substance in equilibrium is zero. One of these various versions is called the *third law of thermodynamics*. Notice that the law was established for $T$ not significantly lower than 1 K.

Statistical mechanically, whether the third law holds or not depends on the energy spectrum of the system; it is easy to make a counterexample for the third law. There is a general belief, however, that the ground state of a real quantum system is not degenerate (due to various small interactions). Still, as is already noted in the text, this is not directly related to the third law.

The third law implies unattainability of $T = 0$.[111]

---

[108] Theodore William Richards (1868-1928) was the first American scientist to receive the Nobel Prize in Chemistry. He demonstrated the existence of isotopes.

[109] [1906: K. Gödel was born on April 28.]

[110] Walter Hermann Nernst (1864-1941)

[111] The third law asserts that the isotherm $T = 0$ coincides with the adiabat $S = 0$. Nernst pointed out that any quasistatic process can be approximated as accurately as one wishes by a combination of isothermal and adiabatic processes, and that only adiabatic processes can reduce the system temperature. Then, he concluded that since there is no other adiabat that can cross $S = 0$, it is impossible to reach $T = 0$ with a finite number of processes. However, the converse does not hold for classical ideal gas. For classical ideal gas, the $T = 0$ isotherms are $P = 0$ and $V = 0$, but any adiabatic process must satisfy $PV^\gamma =$ constant, where $\gamma$ is the ratio of constant pressure and constant volume specific heats: $C_P/C_V$. Thus, $T = 0$ is unattainable for classical ideal gas, but its entropy in the $T = 0$ limit is not bounded from below.

## 1.5 Simple examples

The reader must have already been tired of simple rudimentary topics of statistical mechanics discussed in rather boring undergraduate courses. For completeness sake, however, let us study three simple systems: classical ideal gas, a noninteracting spin system, and a freely jointed polymer. Each simple example tells us some interesting mathematics or physics facts worth remembering: nature of high dimensionality, negative temperature, and adiabatic cooling.

**Classical ideal gas$_{\mathbf{r}}$**

Let us consider the simplest macroscopic system. It is a system consisting of $N$ identical point particles of mass $m$ without any internal structure (degree of freedom) confined in a 3D cube of $V = L \times L \times L$. The box here is understood as an infinite depth (3D) cubic potential well.[112] Since there is no particle-particle interaction, we have only to consider the single-particle Schrödinger equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi = E\psi \tag{1.5.1}$$

with $\psi = 0$ at the box walls. The energy eigenvalues are specified as $E = h^2\boldsymbol{n}^2/(8mL)_{\mathbf{q}}$ by a three-dimensional vector whose components are positive integers $\boldsymbol{n} = (l, m, n)$. The corresponding eigenstate is

$$\psi_{\boldsymbol{n}}(\boldsymbol{r}) = \left(\frac{2}{L}\right)^{3/2} \sin\frac{l\pi x}{L} \sin\frac{m\pi y}{L} \sin\frac{n\pi z}{L}. \tag{1.5.2}$$

In practice, we only need the energy formula, which may be obtained from the general formula for the wavelength of the de Broglie stationary wave with nodes at the walls $\lambda = 2L/n$ ($n$ is a positive integer), the momentum-wavelength relation $p = h/\lambda$, and the formula for the kinetic energy $K = p^2/(2m)$, appropriately generalized to a three dimensional space.

Each eigenstate of particle $i$ is specified by three quantum numbers taking positive integers $\boldsymbol{n}_i = \{l_i, m_i, n_i\}$. The microstate of this system is completely specified by a set of $N$ positive integer three-dimensional vectors (3-vectors) $\{\boldsymbol{n}_1, \cdots, \boldsymbol{n}_i, \cdots, \boldsymbol{n}_N\} = \{\boldsymbol{n}_i\}_{i=1}^N$. The system is isolated, so its total energy is confined in a well-defined energy

---

[112] Although particles do not interact with each other, we assume that each particle is distributed uniformly in all one particle state in the potential well.

shell $\omega_E$ between $E - \delta E$ and $E$. That is,

$$\sum_{i=1}^{N} \frac{h^2 \boldsymbol{n}_i^2}{8mL^2} \in \omega_E. \tag{1.5.3}$$

Therefore, to obtain the number of microstates in this energy shell, we have to count the number of $3N$-dimensional vectors $(\cdots, \boldsymbol{n}_i, \cdots)$ satisfying (1.5.3). That is, we have only to be able to count the number of the integer vectors in the $3N$-dimensional spherical shell of radius $\sqrt{8mE}L/h$ with thickness proportional to $\delta E$. As we will see just below we need not worry about the choice of $\delta E$ very much.

We are studying a very large system, so we should exploit the magnitude of $N$ (high dimensionality of the state space). Calculation of the volume of $\omega_E$ is essentially to obtain the volume of a thin skin of a large dimensional ball.[113] The most important fact to remember about a high dimensional object is that its most volume is in its skin.**q** It is easy to see this. Let the linear dimension of a $D$-object[114] be $R$ and its volume be $CR^D$, where $C$ is a constant dependent on the shape. We skin it. If the skin thickness is $\delta R \ll R$, then the skinned object is similar to the original shape, and its linear dimension is $R - \delta R$. Thus, the ratio of the skinned volume to the original volume is

$$\frac{C(R - \delta R)^D}{CR^D} = \left(1 - \frac{\delta R}{R}\right)^D. \tag{1.5.4}$$

Therefore, for example, even if $\delta R/R = 10^{-10}$, if $D = 10^{20}$, this ratio is almost zero (that is, if $1 \gg \delta R/R \gg D^{-1}$, the volume is almost concentrated in the skin). For high-dimensional objects, its volumes is almost in its skin, and this situation does not change unless the skin is excessively thin. Thus, $\delta E/E$ may be chosen to be very small if not too small. In the actual cases, the ratio is a small number independent of the number of particles, because it is natural to regard $\delta E$ as a leeway in an extensive quantity.

The above consideration implies that the microcanonical partition function $w(E, \{X_i\})$ may be replaced by the number of all the microstates whose energy is less than or equal to $E$ (but all the work coordinates $\{X_i\}$ are confined in the small ranges as

---

[113] We use the standard mathematical terminology; 'sphere' is the surface of a 'ball,' which is a solid object.

[114] Following the tradition of mathematics an object whose dimension is $D$ may be denoted as $D$-object. Three-dimensional space is 3-space. An ordinary ring is 1-sphere. The surface of 3-ball is a 2-sphere, etc.

$w$):

$$W(E, \{X_i\}) = \sum_{0 \le E' \le E} w(E', \{X_i\}) \simeq w(E, \{X_i\}), \tag{1.5.5}$$

where the summation is over all the energy shells up to energy $E$. Thus, we may write Boltzmann's principle as

$$S = k_B \log W(E, \{X_i\}). \tag{1.5.6}$$

Analogous formulas for classical systems also hold.

The total number of states $W(E, V)$ of $N$ point particles in a volume $V$ whose total energy is no more than $E$ can be written as

$$W(E, V) = \frac{1}{N!} \sum_{\{\boldsymbol{n}_i\}: \sum \boldsymbol{n}_i^2 \le 8mEL^2/h^2} 1. \tag{1.5.7}$$

Here, $1/N!$ is put to take into account the indistinguishability of particles as we discussed in the preceding section. The sum may be approximated by the volume of the positive coordinate sector (i.e., the portion in which all the coordinates are positive) of a $3N$-ball of radius $\sqrt{8mE}L/h$. Therefore,

$$W(E, V) = \frac{1}{N!} \frac{1}{2^{3N}} B_{3N}(\sqrt{8mE}L/h), \tag{1.5.8}$$

where $B_D(R)$ is the volume of $D$-ball of radius $R$. It has the general form $K_D R^D$, where $K_D$ is a constant, which we will compute later, but for most purposes we do not need any explicit form.

From (1.5.8) we see

$$W(E, V) \propto E^{3N/2} L^{3N} = E^{3N/2} V^N, \tag{1.5.9}$$

which may also be obtained easily by dimensional analysis.[115] **q** Let us first see that we can obtain important thermodynamic relations, using this dimensional analytic result and a general consideration (the extensivity of entropy, etc.). Boltzmann's principle tells us

$$S = k_B N \left\{ \log V + \frac{3}{2} \log E + c_N \right\}. \tag{1.5.10}$$

---

[115] Quantum mechanically, we must make a dimensionless quantity from $V$ and $E$. We may use $m$ and $h$. Then, since $h/\sqrt{2mE}$ has a dimension of length, we have a dimensionless quantity $(\sqrt{2mE}/h)^3 V$.

Here, since the proportionality constant in (1.5.9) depends on $N$, the remainder $c_N$ is explicitly written with suffix $N$. We know $S$, $V$ and $E$ must be extensive (proportional to $N$), so we see

$$S = k_B N \left\{ \log \frac{V}{N} + \frac{3}{2} \log \frac{E}{N} + c \right\}. \tag{1.5.11}$$

Here $c$ is a constant which should not depend on $N$ (if $N$ is sufficiently large). From this with the aid of thermodynamics we immediately obtain $PV = Nk_BT$ and $E = 3Nk_BT/2$. The latter implies that each translational degree of freedom has the average energy $k_BT/2$ (equipartition of energy, see Section 2.2). From these relations we get $PV = 2E/3$.[r] The relation is, as we will see in Chapter 3, valid for quantum ideal gases as well. It is easy to see from (1.5.10) that this relation is directly obtained from $P = -(\partial E/\partial V)_S$. This fact suggests that the relation $PV = 2E/3$ is a more mechanical than a thermal relation. It is indeed the case, as we can derive this relation immediately from the virial theorem as we will see in Chapter 4. The 2 in this formula is because the energy is the quadratic form of momenta, and the 3 is the spatial dimensionality. If we notice such relations, it is easy to generalize the relation to arbitrary dimensional spaces and to superrelativistic gases.

To calculate $c$ in (1.5.10) we must calculate the volume of the $D$-ball. Here, we give supposedly the cleverest method.

**The volume of a $D$-ball**[r]
(1) Let $S_D$ be the surface area of the unit $(D-1)$-sphere (the boundary of a unit $D$-ball). Show that $B_D(R) = (S_D/D)R^D$.
(2) Demonstrate that

$$I_D = \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_D e^{-a(x_1^2 + \cdots + x_D^2)} = (\pi/a)^{D/2}, \tag{1.5.12}$$

where $a$ is a positive constant. Compute the same integral in the polar coordinate system to show (cf. in 2-space we use $2\pi r dr$, in 3-space $4\pi r^2 dr$)

$$I_D = \frac{1}{2} S_D \Gamma(1 + D/2) a^{-D/2}. \tag{1.5.13}$$

(3) Comparing (1.5.12) and (1.5.13), we obtain[q]

$$S_D = \frac{2\pi^{D/2}}{\Gamma(D/2)}. \tag{1.5.14}$$

Thus, we have obtained

$$B_D(R) = \frac{\pi^{D/2}}{\Gamma(D/2 + 1)} R^D. \tag{1.5.15}$$

From (1.5.15) we obtain

$$W(E,V) = \frac{V^N}{N!h^{3N}} \frac{(2\pi mE)^{3N/2}}{\Gamma(3N/2+1)}. \tag{1.5.16}$$

That is, with the aid of Stirling's formula $\Gamma(N) \simeq N! \simeq (N/e)^N$ for $N \gg 1$,

$$S(E,V,N) = Nk_B \left\{ \log\frac{V}{N} + \frac{3}{2}\log\frac{E}{N} + \log\frac{(4\pi m/3)^{3/2}e^{5/2}}{h^3} \right\}. \tag{1.5.17}$$

Using $E = 3Nk_BT/2$, we can rewrite this as

$$S(E,V,N) = Nk_B \left\{ \log\frac{1}{n} + \frac{3}{2}\log\frac{mk_BT}{2\pi\hbar^2} + \frac{5}{2} \right\}, \tag{1.5.18}$$

where $n = N/V$ is the number density. The de Broglie thermal wave length[r] $\lambda_T = (h^2/mk_BT)^{1/2}$ is obtained from the average kinetic energy $\langle p_x^2\rangle/2m = (1/2)k_BT$ and de Broglie's relation between the wavelength of the matter wave and momentum $\lambda = h/p$. The number density $n_Q$[r] of the ideal gas that contains on the average 1 particle in the cube with the edge length $\lambda_T/\sqrt{2\pi}$, i.e.,[q]

$$n_Q = (mk_BT/2\pi\hbar^2)^{3/2} = (\sqrt{2\pi}/\lambda_T)^3 \tag{1.5.19}$$

is called the *quantum number density*. Then, (1.5.18) may be expressed as

$$S = k_BN[\log(n_Q/n) + 5/2]. \tag{1.5.20}$$

Numerically, $n_Q$ may be computed as[q]

$$n_Q = 9.88 \times 10^{29}\hat{m}^{3/2}\left(\frac{T}{300}\right)^{3/2} \simeq \hat{m}^{3/2}\left(\frac{T}{300}\right)^{3/2} \times 10^{30}, \tag{1.5.21}$$

where $\hat{m}$ is the mass of the particle in atomic mass unit (e.g., for water $\hat{m} = 18$). Since $n_Q$ goes to zero in the $T \to 0$ limit, $S$ is not bounded from below, contradicting the third law of thermodynamics. It is not a desirable (self-contradicting) practice to use the Carnot cycle with the aid of the ideal gas law to discuss the second law of thermodynamics. In any case, it is a bad taste to use a particular example to demonstrate general statements.

The chemical potential can be obtained as[q]

$$\mu = k_BT\log(n/n_Q) \tag{1.5.22}$$

from (1.5.17) ((1.5.20) is not convenient, since $E$ is not explicit) . If we replace $n_Q$ with an appropriate effective number density $n^{\ominus}$ at a standard state that may depend on $T$, $P$, etc., this form of chemical potential can be used for any non-interacting (classical) particles, e.g., solute molecules in dilute solutions, rare excitations in solids, etc. This chemical potential becomes $-\infty$ in the $n \to 0$ limit.

**Dilution limit**

Suppose a chemical reaction changes a chemical $A$ whose concentration is $n_A$ into $B$ whose concentration is 0. Then, this chemical reaction produces 'infinite' free energy that can be converted into work. Thermodynamically nothing is wrong, but this sounds too good. What happens, actually?

It is said that the 'grade' of chemical energy is lower than that of mechanical energy, and that of heat energy is further lower than that of chemical energy. What does this refer to?

If we wish to do statistical mechanics within classical mechanics, there is no concept of the number of states $W(E,V)$, since there is no concept of energy level, but statistical mechanics existed before quantum mechanics. Instead of $W(E,V)$, a quantity proportional to it, i.e., the volume $\tilde{W}(E,V)$ of the phase space with energy less than $E$:

$$\tilde{W}(E,V) = \frac{1}{N!}\int_{\boldsymbol{q}_1\in V,\cdots,\boldsymbol{q}_N\in V,\,(\boldsymbol{p}_1^2+\cdots+\boldsymbol{p}_N^2)/2m\le E} d\Gamma \tag{1.5.23}$$

was used with the Boltzmann's famous formula, where $d\Gamma = d^3\boldsymbol{q}_1\cdots d^3\boldsymbol{q}_N d^3\boldsymbol{p}_1\cdots d^3\boldsymbol{p}_N$ is the phase volume element. Here, the factor $1/N!$ coming from the indistinguishability of identical particles is taken into account. This can be computed in terms of the high-dimensional sphere volume $B_{3N}$ as

$$\tilde{W}(E,V) = \frac{1}{N!}V^N B_{3N}(\sqrt{2mE}). \tag{1.5.24}$$

Since $B_{3N}(ar) = a^{3N}B_{3N}(r)$, we obtain from (1.5.8)

$$W(E,V) = \frac{1}{h^{3N}}\tilde{W}(E,V) \tag{1.5.25}$$

Thus, it is sensible to relate the phase integral and quantum mechanical trace over all the states as

$$Tr \to \frac{1}{N!h^{3N}}\int d\Gamma. \tag{1.5.26}$$

In other words, $h^{3N}$ may be understood as the phase volume occupied by a single microstate.[116]

### Non-interacting spin system in an external magnetic field$_{\mathbf{q}}$

If magnetic atoms are dilute in an insulating solid, they do not interact with each other;[117] such a dilute magnet may be understood as a collection of non-interacting spins. Here, for simplicity, the spin value is ±1 and its magnetic moment is $\mu$ times the value of the spin.

In the magnetic field $B$ ($\neq 0$, in $z$ direction), the system Hamiltonian is usually written as

$$H = -\sum_i \mu B \sigma_i, \tag{1.5.27}$$

where $\sigma_i$ is the $i$th spin ($\pm 1$ for the $\pm z$ direction, respectively). Does this Hamiltonian express the energy of the system? Since there is no interaction among spins, this (rather artificial) system should not have internal energy (or always $E = 0$[118]). This is indeed true. This Hamiltonian (1.5.27) describes the potential energy of spins placed in a magnetic field. In other words, it is the interaction Hamiltonian between the spin system and the system M creating the external magnetic field $B$ (or you could say the interaction energy between the spin system and the magnetic field $B$). Therefore, it is natural to consider the system as a thermally isolated system whose Hamiltonian is given by (1.5.27). The natural thermodynamic potential is $J = E - BM$, where $M$ is magnetization.[119] Our system is quite artificial, and always $E = 0$, so $J = -BM$. The natural thermodynamic variables are $J$ and $B$, so

---

[116] The above argument is exact only for ideal gas, so many readers would say that it is not generally convincing. There is a formal argument with the aid of the Wigner representation. For this formal argument, see Supplementary Pages.

[117] ⟪**Ordering at very low temperatures**⟫ If there is no interaction, the system entropy stays very large even at $T = 0$, so such a system contradicts thermodynamics. In reality, however dilute the spin system is, there are residual interactions among spins, which cannot be ignored at sufficiently low temperatures. Furthermore, although here magnetic dipole interactions are ignored, they are not only long-ranged but also without any shielding effect, so very serious problems occur. For example, the system thermodynamic functions depend on the sample shape. Here, we consider only the situations where we may ignore the spin-spin interactions.

[118] In the actual system, spins are riding on atoms, so we cannot ignore, e.g., kinetic energy of atoms as we will discuss later, but for the time being, we assume as if there are only spins.

[119] $J$ may be called a generalized enthalpy. Compare $dE = TdS + BdM$ in the present case and that for a gas system: $dE = TdS - PdV$. The usual enthalpy is $H = E - (-PV) = E + PV$.

we should count $w(J,B)$, the number of eigenstates of (1.5.27) in the shell $(J-\delta J, J]$.

Let $N_+$ be the number of upward spins and $N_-$ that of downward spins. Then, the magnetization $M$ reads $M = \mu(N_+ - N_-)$, so the number of microstates consistent with the macrostate $(-BM, B)$ is given by

$$w(J,B) = \frac{N!}{N_+!N_-!}. \tag{1.5.28}$$

Here, note that $N_\pm = (N \pm M/\mu)/2 = (N \mp J/\mu B)/2$. Using Boltzmann's principle (translation rule) with the aid of Stirling's formula $N! \sim (N/e)^N$, we have

$$S = -k_B(N_+ \log N_+ + N_- \log N_- - N \log N). \tag{1.5.29}$$

The Gibbs relation is $dJ = TdS - MdB$. Consequently, we obtain

$$\frac{1}{T} = \left.\frac{\partial S}{\partial J}\right|_B = \frac{k_B}{2\mu B}\log\frac{N_+}{N_-}, \tag{1.5.30}$$

or

$$M = \mu N \tanh \beta\mu B, \tag{1.5.31}$$

where $\beta = 1/k_B T$ is a standard notation.[120]

Starting from a low temperature, if we increase the temperature, $N_-/N_+ \nearrow 1$. This ratio is at most one even in the $T \to \infty$ limit. However, if we flip all the spins in the magnetic field, $N_- > N_+$ is realized easily,[121] and it corresponds to $\beta < 0$ as can be seen from (1.5.31). This is a *negative temperature* state. This is a higher energy state than any positive temperature state, so we may conclude that the negative temperature is actually 'hotter' than $T = +\infty$. This sounds natural in terms of $\beta$, since $\beta \searrow$ implies higher temperature.

Let us 'measure' the temperature of a negative temperature state with a gas thermometer (e.g., a monatomic ideal gas with $n$ particles, so $P \propto T$ and $E =$

---

[120] The idealized spin system we are discussing here is simple but an over-idealized artificial system. It is thermodynamically strange. For example, the internal energy is always $E = 0$ (or constant) no matter what happens to spin configurations, because there is no spin-spin interaction at all; even if entropy changes, energy remains unchanged. To change entropy, the magnetization must be changed, so the standard thermodynamic relation $(\partial E/\partial S)_M = T$ is meaningless, so no contradiction actually happens. That is, the system is indeed thermodynamically strange, but it does not contradict thermodynamics in contrast to classical ideal gas.

[121] Here, the up or down direction of the spin is defined relative to the magnetic field as above. To flip spins we can use an electromagnetic pulse, which is an important experimental means in magnetic resonance.

$3nk_BT/2$).[122] Let the negative temperature state we wish to observe have magnetization $M_0 = \mu(N_+ - N_-) < 0$ ($E = B|M_0|$), and the state of the spin system after equilibrating with the thermometer be with magnetization $M > 0$ (it must be a positive temperature state). The total energy including the potential energy $-hM$ is conserved, so

$$B|M_0| + \frac{3}{2}nk_BT_0 = -MB + \frac{3}{2}nk_BT. \tag{1.5.32}$$

Because $|M_0| = O[N]$, we must conclude that $3nk_BT/2$ must also be of order $N$. That is $T = O[N/n]$. A thermometer must be a tiny system, so this ratio must be huge: $T \gg 1$. Then, $M$ is almost zero, so

$$T = \frac{2B|M_0|}{3k_Bn} \tag{1.5.33}$$

must be accurate. The negative temperature state, though destroyed upon contacting with the thermometer, was indeed very hot. If the reader wishes to solve this problem more 'mechanically' and to confirm this is indeed correct, use (1.5.31) and then solve (1.5.32).

This calculation explains why negative temperature systems cannot exist for a long time (i.e., as a stable equilibrium state), because spin systems are always coupled with a crystal lattice that has no upper energy bound.[123]

In reality,**r** each spin interacts with the degrees of freedom of the lattice in which it is sitting. If we include the contribution of this interaction, the entropy of the system should have the following form

$$S = S(M) + S_e(E, M). \tag{1.5.34}$$

$S_e(E, M)$ reflects the effect of the lattice system whose number of states increases rapidly with the increase of energy, so due to the presence of $S_e$, the 'real' ideal spin system cannot have negative temperature states in equilibrium. The internal energy is no more constant even if $B = 0$, or $S$ is no more constant even if $M$ is constant, so $E$ and $M$ are independent thermodynamic coordinates and the Gibbs relation can tell us a nontrivial relation[124]

$$dE = TdS + BdM. \tag{1.5.35}$$

---

[122] This is only a gedanken experiment. As can be seen easily, the temperature is so high that even relativistic corrections may be needed. Here, we proceed naively without such corrections.

[123] Precisely speaking, even if the energy is not bounded from above, negative temperature states are realizable as stable states, if the number of microstates does not increase sufficiently with energy.

[124] In the over-idealized free spin case, $dE = TdS + BdM$ is still correct, but since $E = 0$, so this gives $dS = -(B/T)dM$, which is of course correct.

Now, let us compare (1.5.34) and the formula for the ideal gas entropy, which is divided into the part dependent solely on the gas-specific variable $V$ and the rest that depends only on internal energy:

$$S = S(V) + S_e(E). \tag{1.5.36}$$

This separation is the true meaning of the adjective 'ideal.' This suggest that the 'true ideal spin system' is the spin system for which $S_e$ in (1.5.34) is dependent only on $E$.

### Freely jointed chain (ideal polymer chain)$_{\mathbf{r}}$

Suppose noninteracting spins are sitting on a one-dimensional lattice (Fig. 1.5.1 lower left). Next, let us regard each arrow as a small molecule (monomer) (an up or down spin corresponds to an up or down monomer, respectively) and connect (polymerize) them in the lattice order. We have a polymer chain in one-dimensional space (Fig. 1.5.1 right).[125] Corresponding to the lack of any interaction among spins, the monomers can freely change its directions (up or down) except for being connected.

Let us assume that the chain starts from the origin, and the number of steps = monomers is $N$, which is called the *degree of polymerization*. The length of a single

---

[125] **⟪Usefulness of spin-monomer correspondence⟫** The reader might think it is unnatural that spins on the lattice do not interact each other, so the 'magnet-polymer' correspondence described here is highly artificial. In reality, there are many interactions among monomers, so we can think them as spin-spin interactions. Actually, this correspondence provides an effective way to compute the properties of polymers (e.g., polypeptides). See P. J. Flory, *Statistical mechanics of chain molecules* (Interscience Publishers, 1969).

monomer is $\ell$.

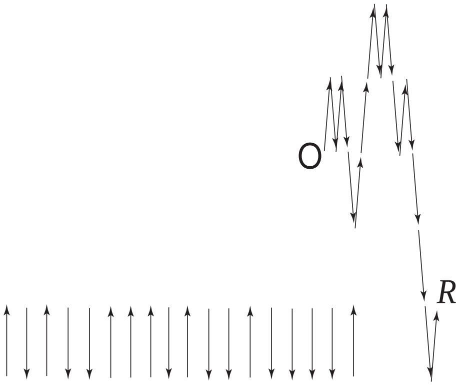


Fig. 1.5.1 A system consisting of independent spins that can assume up or down state may be identified with a polymer chain. The end-to-end 'vector' $\boldsymbol{R}$ corresponds to the magnetization. Here, to make the figure clear spins = steps are tilted slightly in the polymer chain.

The translation of spin system quantities into the polymer chain quantities is: $M \to L$, $h \to F$, and $\mu \to \ell$, where $L$ is the end-to-end distance, and $F$ the external tensile force.[126] Since $N_\pm = (N \pm L/\ell)/2$, the result (1.5.30) for the spin system tells us

$$F = \frac{1}{2} k_B T \log \frac{N + L/\ell}{N - L/\ell}, \tag{1.5.37}$$

or

$$L = N\ell \tanh \frac{\ell F}{k_B T}. \tag{1.5.38}$$

We see that the chain end-to-end distance (polymer length) shrinks upon heating under a constant external force. This is the characteristic feature of *entropic elasticity*—elasticity due to entropy change in contrast to the elasticity of solid due mainly to energy change— that is the essence of rubber elasticity.

Just as the spin system discussed above, this polymer system is also a rather artificial system without any interactions among monomers. Furthermore, kinetic energy of monomers is totally neglected, so the internal energy $E$ of the polymer is always constant (zero) independent of $T$. Its thermodynamic space is virtually

---

[126] The Hamiltonian of a stretched polymer under the influence of $F$ is often written as

$$H = -\sum_i F\ell\sigma_i,$$

by mimicking the spin system, $\sigma_i = \pm 1$ ($+1/-1$ is upward/downward), but this is not the Hamiltonian of the polymer system, but the interaction Hamiltonian between the polymer and the system producing force $F$ (just as noted for the 'ideal spin' case (1.5.27)). That is, this $H$ is the potential energy of the polymer in the force field. There is a minus sign, because to shorten the chain against the external force requires energy.

spanned by $L$ only, and $S$ cannot be changed independently from $L$. Therefore, the model is not enough to discuss real polymer chains. In reality, monomers undergo thermal fluctuation, but its contribution is totally ignored in this model. To consider the contribution of thermal fluctuation to entropy, we must add a portion $S_e(E, L)$ that depends much more on internal energy than on conformations (or $L$) as[127]

$$S = -k_B \left( N_+ \log \frac{N_+}{N} + N_- \log \frac{N_-}{N} \right) + S_e(E, L). \tag{1.5.39}$$

For such a model the Gibbs relation $dE = TdS + FdL$ can give nontrivial relations.

The standard statistical mechanics approach may be to make a more realistic microscopic model and to compute a more realistic entropy formula, but general qualitative features may be obtained thermodynamically in a more robust fashion. Therefore, let us pursue logical consequences of the characteristic feature of entropic elasticity: if the chain is warmed up under a constant force, it shrinks. That is,

$$\left. \frac{\partial L}{\partial T} \right|_F < 0. \tag{1.5.40}$$

It is a good occasion to brush up the reader's elementary skill. Also she may be able to recognize the power of thermodynamics.[128]

With the aid of a Maxwell's relation[129] we can rewrite (1.5.40) as

$$\left. \frac{\partial L}{\partial T} \right|_F = \left. \frac{\partial S}{\partial F} \right|_T = \left. \frac{\partial S}{\partial L} \right|_T \left. \frac{\partial L}{\partial F} \right|_T < 0. \tag{1.5.41}$$

---

[127] If $S_e(E, L)$ does not depend on $L$ at all, the chain is called an ideal polymer chain. Recall ideal gas and ideal spin systems.

[128] The contraction of a loaded rubber band upon heating was discovered by J. Gough in 1805. Joule discovered that adiabatic stretching generates heat in 1857. This was the first observation of adiabatic change in nongaseous substance.

[129] Our basic Gibbs relation is $dE = TdS + FdL$, but the independent variables of (1.5.40) are $T$ and $F$, so we consider the following Legendre transformation $E - ST - FL$, which gives the transformed Gibbs relation containing $-SdT - LdF$.

As we will see in [T9], we use Maxwell's relation in the form of $\partial(x, X)/\partial(Y, y) = 1$ in this book. Here, the upper and lower case letter pair of the same alphabet implies a conjugate extensive-intensive variable pair (with respect to energy). The calculation in the present example looks like:

$$\left. \frac{\partial L}{\partial T} \right|_F = \frac{\partial(L, F)}{\partial(T, F)} = \frac{\partial(L, F)}{\partial(T, S)} \frac{\partial(T, S)}{\partial(T, F)} = \frac{\partial(T, S)}{\partial(T, F)} = \left. \frac{\partial S}{\partial F} \right|_T .$$

Thus, since $(\partial L/\partial F)_T > 0$, if the chain is stretched, the entropy of the chain decreases under constant temperature:

$$\left.\frac{\partial S}{\partial L}\right|_T < 0. \tag{1.5.42}$$

This can be intuitively visualized as Fig. 1.5.2.

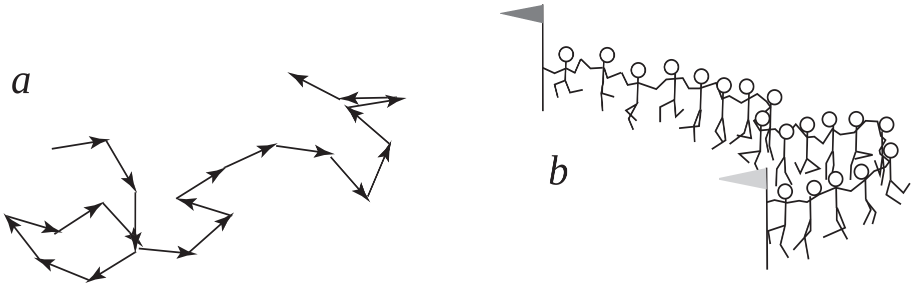


Fig. 1.5.2 The temperature represents how vigorously kids are moving around. This also includes 'vibration' of individual bodies. The figure is after N. Saito, *Polymer Physics* (Shokabo, 1967) (The original picture was due to T. Sakai's lecture according to Saito). Entropy is monotonically related to the width of the range kids can play around easily, which becomes smaller if the distance between the flags is increased.

How can we check (1.5.40) with an easy experiment? The length change upon warming up is not very large.[130] We should observe a phenomenon more easily observed and logically connected to the length change. If the chain is adiabatically stretched, then the kids in Fig. 1.5.2 cannot run around easily, but the entropy must be maintained, so they must keep 'moving.' That is, they must move their bodies vigorously without running around to make up the entropy lost due to the restricted running range. Thus, we guess the temperature goes up; we expect

$$\left.\frac{\partial T}{\partial L}\right|_S > 0. \tag{1.5.43}$$

During a reversible adiabatic process, entropy is constant. Since thermodynamics can compute derivatives for reversible changes only, a derivative under adiabatic condition is always the derivative under constant entropy. Let us first check the

---

[130] However, still it is possible to observe the length change with the naked eye. Strongly stretch a rubber band with a mark around its center and fix its both ends. Then, heat, say the right half of the rubber band with a hair dryer. We can observe that the central mark is pulled to the right slightly.

consistency of (1.5.42) and (1.5.43) (for an explanation of Jacobian technique used here, see the fine-lettered explanation [T9]):

$$\left.\frac{\partial S}{\partial L}\right|_T = \frac{\partial(S,T)}{\partial(L,S)}\frac{\partial(L,S)}{\partial(L,T)} = -\frac{C_L}{T}\left.\frac{\partial T}{\partial L}\right|_S. \tag{1.5.44}$$

Here, $C_L$ is the heat capacity of the polymer chain under constant length.[131] The heat capacity is positive, so the partial derivative in (1.5.42) and that in (1.5.43) must have the opposite signs. (1.5.43) can be easily experienced by stretching a small portion of a wide rubber band (such as used to bundle broccoli) rapidly and put the part being kept stretched to the lip; the reader will feel warmth.[132] Can the reader guess the sign of

$$\left.\frac{\partial S}{\partial F}\right|_L ? \tag{1.5.45}$$

**[T9] Jacobian technique**

To manipulate many partial derivatives, it is very convenient to use the so-called *Jacobian technique*. The Jacobian of two functions $X$ and $Y$ of two independent variables $x$, $y$ is defined as the following determinant:

$$\frac{\partial(X,Y)}{\partial(x,y)} \equiv \begin{vmatrix} \left.\frac{\partial X}{\partial x}\right|_y & \left.\frac{\partial X}{\partial y}\right|_x \\ \left.\frac{\partial Y}{\partial x}\right|_y & \left.\frac{\partial Y}{\partial y}\right|_x \end{vmatrix} = \left.\frac{\partial X}{\partial x}\right|_y \left.\frac{\partial Y}{\partial y}\right|_x - \left.\frac{\partial Y}{\partial x}\right|_y \left.\frac{\partial X}{\partial y}\right|_x. \tag{1.5.46}$$

In particular, we have

$$\frac{\partial(X,y)}{\partial(x,y)} = \left.\frac{\partial X}{\partial x}\right|_y. \tag{1.5.47}$$

It should not be hard to guess the cases with more variables and functions. For example, when we must keep many variables constant

$$\left.\frac{\partial X}{\partial x}\right|_{x_1,\cdots,x_n} = \frac{\partial(X,x_1,\cdots,x_n)}{\partial(x,x_1,\cdots,x_n)} \tag{1.5.48}$$

From the property of the determinant, if we change the order of variables or functions, there is a sign change (this is true for general cases larger than 2 × 2):

$$\frac{\partial(X,Y)}{\partial(x,y)} = -\frac{\partial(X,Y)}{\partial(y,x)} = \frac{\partial(Y,X)}{\partial(y,x)} = -\frac{\partial(Y,X)}{\partial(x,y)}. \tag{1.5.49}$$

---

[131] ⟪**Heat capacity and entropy**⟫ Suppose the system temperature increases by $\Delta T$ when heat $\Delta Q$ is added to it under constant $X$. Then, $\Delta Q/\Delta T = T\Delta S/\Delta T$ must be equal to the heat capacity $C_X$. This is a method to measure the heat capacity.

[132] We use our lips, because it is sensitive to temperature. Use rubber bands that are quite unlikely to snap lest you have very painful experience.

If we assume that $X$ and $Y$ are functions of $a$ and $b$, and that $a$ and $b$ are, in turn, functions of $x$ and $y$, we have the following multiplicative relation:

$$\frac{\partial(X,Y)}{\partial(a,b)}\frac{\partial(a,b)}{\partial(x,y)} = \frac{\partial(X,Y)}{\partial(x,y)}. \tag{1.5.50}$$

This is a disguised chain rule. The proof of this relation is left to the reader. Use

$$\left.\frac{\partial X}{\partial x}\right|_y = \left.\frac{\partial X}{\partial a}\right|_b \left.\frac{\partial a}{\partial x}\right|_y + \left.\frac{\partial X}{\partial b}\right|_a \left.\frac{\partial b}{\partial x}\right|_y . \tag{1.5.51}$$

The rest is straightforward algebra. From (1.5.50) we get at once

$$\frac{\partial(X,Y)}{\partial(x,y)} = 1 \Big/ \frac{\partial(x,y)}{\partial(X,Y)} . \tag{1.5.52}$$

In particular, we have

$$\left.\frac{\partial X}{\partial x}\right|_Y = 1 \Big/ \left.\frac{\partial x}{\partial X}\right|_Y . \tag{1.5.53}$$

Using these relations, we can easily demonstrate

$$\left.\frac{\partial X}{\partial y}\right|_x = - \left.\frac{\partial x}{\partial y}\right|_X \Big/ \left.\frac{\partial x}{\partial X}\right|_y \tag{1.5.54}$$

as follows:

$$\frac{\partial(X,x)}{\partial(y,x)} \overset{(1.5.50)}{=} \frac{\partial(y,X)}{\partial(y,x)}\frac{\partial(X,x)}{\partial(y,X)} \overset{(1.5.49)}{=} -\frac{\partial(x,X)}{\partial(y,X)}\frac{\partial(X,y)}{\partial(x,y)}. \tag{1.5.55}$$

Then, use (1.5.52). A concrete example of this formula is

$$\left.\frac{\partial P}{\partial T}\right|_V = - \left.\frac{\partial V}{\partial T}\right|_P \Big/ \left.\frac{\partial V}{\partial P}\right|_T , \tag{1.5.56}$$

which relates thermal expansivity and isothermal compressibility.

Let $\{x, X\}$ and $\{y, Y\}$ be two conjugate pairs of thermodynamic variables with respect to energy. Notice that all the Maxwell's relations can be unified as

$$\frac{\partial(x,X)}{\partial(y,Y)} + 1 = 0 \quad \text{or} \quad \frac{\partial(x,X)}{\partial(Y,y)} = 1. \tag{1.5.57}$$

Confirm the statement.[133] When this formula is used, do not forget that the conjugate

---

[133] When the author mentioned this to his colleague, M. Stone, he showed me: V. Ambegaokar and N. D. Mermin, "Question # 78. A question about the Maxwell relations in thermodynamis," Am. J. Phys., **69**, 405 (2001), which asks, 'where in the literature this ($= \partial(P,V)/\partial(T,S) = 1$) pleasing and apparently little known way of looking at Maxwell's relations can be found?" The answer appeared in Am. J. Phys., **70**, 105 (2002) by D. J. Ritchie and by H. S. Leff. The former wrote actually Mermine taught this in 1967 [a demonstration is given in D. J. Ritchie, Am. J. Ritchie, "A simple method for deriving Maxwell's relations," Am. J. Phys., **36**, 760 (1968)] The latter quotes this same paper. The derivation given is lengthy, but this is almost a trivial result of the differential forms: $d^2E = 0$.

variable of $V$ is $-P$ (do not forget the sign). Needless to say, we can use

$$\frac{\partial(-X,Y)}{\partial(x,y)} = -\frac{\partial(X,Y)}{\partial(x,y)} = \frac{\partial(X,Y)}{\partial(-x,y)}. \tag{1.5.58}$$

(1.5.43) also implies

$$\left.\frac{\partial T}{\partial L}\right|_S = \left.\frac{\partial T}{\partial F}\right|_S \left.\frac{\partial F}{\partial L}\right|_S > 0 \Rightarrow \left.\frac{\partial T}{\partial F}\right|_S > 0. \tag{1.5.59}$$

That is, if we suddenly[134] remove the tensile force when the rubber band is stretched, its temperature can be lowered. Or, in the magnetic system language

$$\left.\frac{\partial T}{\partial B}\right|_S > 0 \tag{1.5.60}$$

allows us to cool the spin system by removing the magnetic field (and reducing the order).[135] This is the principle of *adiabatic cooling*.

More generally, assume that the system entropy depends on an intensive parameter $\alpha$ that can increase the order in the system. Tensile force $F$ or magnetic field $B$ is such a parameter that can control the alignment of monomers or spins. Suppose that we can increase $\alpha$ adiabatically (and in a quasiequilibrium fashion). Then, $S(T,\alpha) =$ const., but if $T$ were maintained, $S$ should have decreased, so increasing $\alpha$ adiabatically should increase the system temperature to maintain the entropy (recall the rubber band).

Now, cool the system to the original temperature, and then return $\alpha$ to the original value. The system should have been cooled. This is the general principle of the adiabatic cooling. In practice, $\alpha$ is increased isothermally, and then adiabatically $\alpha$ is reduced to the original value. The system cools down as illustrated by the path in Fig. 1.5.3.

---

[134] In practice, we can simply relax a stretched rubber band rapidly. This relaxation is far slower than the molecular relaxation rates, so the resultant process is virtually quasistatic. Needless to say, the warming up due to heat influx from the ambient air is time consuming, so we can realize a reversible adiabatic process.

[135] For a noninteracting spin system, as seen from (1.5.31), $B/T$ is constant under the $S$ constant condition, so we can see this explicitly.

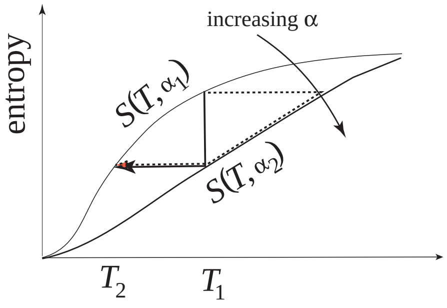


Fig. 1.5.3 Adiabatic cooling.
Initially, the system is at $T_1$. Isothermally, $\alpha$ is increased as $\alpha_1 \to \alpha_2$. This decreases entropy. Now, $\alpha$ is returned to the original smaller value adiabatically. The entropy is maintained, and the temperature decreases (adiabatic cooling) to $T_2$. The dotted path is the one explained initially in the text.

## Problems for Chapter 1

**1.1** [Equivalence of heat and work]
A block of mass $M = 1$ g is at rest in space (vacuum). Another block of the same mass and velocity $V = 1.5$ km/s collides with the first block and the two blocks stick to each other.
(1) Assuming that there is no radiation loss of energy and that there is no rotation of the resultant single block, obtain the temperature of the resultant single block after equilibration. Assume that the specific heat of the material is 2.1 J/g·K.
(2) If rotation of the resultant body is allowed, what can be said about its final temperature? In particular, is it possible not to raise the temperature of the resultant single block? (Only a qualitative discussion will be enough.)

**1.2** [Exchange of temperatures]
Suppose there are two water tanks A and B containing the same amount of water. Initially, A is 42°C and B is 25°C. The final state we want is A to be 25°C and B 42°C (that is, the temperatures of A and B are exchanged; e.g., A is the used warm bath water, and B new clean tap water). Assume that the whole system is thermally isolated.
(1) Is the process reversibly feasible? Explain why.
(2) Devise a process. No precise statement is needed; only state key ideas.
(3) If the reader's process in (2) contains mechanical parts (e.g., a heat engine), devise a process without any moving parts. No precise statement is needed; only state key ideas. The reader can use other containers to hold water and can freely move or flow water (ignore dissipation).

**1.3** [The fourth law of thermodynamics]
(1) For 0.5 moles of a certain substance the equation of state is obtained as:

$$E = \kappa T V^{1/2}, \tag{1.P.1}$$

where $E$ is internal energy, $V$ is volume, $T$ is absolute temperature, and $\kappa$ is a constant. Write down the equation of state for $N$ moles of this substance.
(2) We can define extensive quantities per molecule $x = X/N$, where $X = E, S, V$ and $x = e, s, v$. Write down the Gibbs relation for one mole (or a molecule). That is, express $de$ in terms of extensive quantities per mole (or molecule).

**1-XX** [Elementary thermodynamics]
The following equations of state cannot be realized, because they violate some general rules of thermodynamics. State explicitly the reason why they are impossible. Notations are standard ones (e.g., $E$ is the internal energy, etc.) and $\alpha$ is a positive constant.

(1) $S = \alpha[N^2E/V]^{1/3}$
(2) $S = \alpha V^3/(NE)$
(3) $S = \alpha(NE)^{1/2}e^{-EV/N}$
(4) $S = N\log(V/N)$.

**1-XX** [Elementary Thermodynamics]
(1) A substance undergoes a melting phase transition at temperature $T_m$ under pressure $p$. The latent heat is $\Delta H$(cal/mol). The specific heat (in cal/molK) in the solid phase is $C_S$ and in the liquid phase $C_L$; they are insensitive to temperature. Let $T_0 < T_m < T_1$. Find the entropy difference between the solid at $T_0$ and liquid at $T_1$.
(2) Find the excess entropy the supercooled liquid at $T_0$ has relative to the solid at the same temperature.[136]

**1-xx** [The most general expression of Maxwell's relations]
Demonstrate that all the Maxwell's relations are unified into the following form:

$$\frac{\partial(x, X)}{\partial(y, Y)} + 1 = 0. \tag{1.P.2}$$

Here, $X$ and $Y$ are generic extensive quantities, and $x$ and $y$ are the corresponding conjugate intensive variables (wrt internal energy).

**1.4** [Asymmetric coin and limit theorem]
The law of large numbers applies to an unfair coin as well. Suppose the coin we use has a probability $1/2 + \varepsilon$ to yield a head (H). Otherwise, a tail (T) is yielded. One get \$1 when H shows up and must pay \$1, otherwise.
(1) Write down the generating function $\omega_N(k)$ for the reader's expected gain per one coin-tossing $s_N$ for the length $N$ coin-tossing sequence.
(2) Compute the mean and the variance of $s_N$.
(3) Using the generating function technique, find the density distribution function $f_N$ for the fluctuation of $s_N$ to demonstrate[137] the law of large numbers and the central limit theorem.

**1-XX** [Elementary probability]
Monty problem.

**1-XX** [Uncorrelatedness and independence]
Let $X$ be a Gaussian stochastic variable, and $Y = \pm X$ depending on the coin tossing (e.g., choose + for head). Show that $X$ and $Y$ are uncorrelated, but not statitically

---

[136] Assume that equilibrium thermodynamics applies to metastable states. This assumption is practically always all right.

[137] Need not be mathematical; quite a theoretical physicist's way is OK!

independent.[138]

**1.5** [How to use Chebyshev's inequality]
(1) We wish to know whether a coin is fair or not. To estimate the probability of the head H within ±0.01, how many throwings do you need? Let us tolerate larger errors once in 100 runs. You may assume that the coin is not grossly unfair.
(2) Theoretically, it is known that if the coin is flipped rapidly, the final surface agrees with the original surface with probability 0.51 (for example, if the original surface is H, then with probability about 0.51 H is obtained). To confirm this bias how many throwings do you think are needed?

**1-xx** [Insufficiency of weak law of large numbers] Let $\{X_n\}$ be iid with

$$P(|X| > t) = e/t \log t \tag{1.P.3}$$

for $t > e$. Show that **1.6** [A Monte Carlo method to determine $\pi$]
There is a unit square ($1 \times 1$ square) and in it is inscribed a disk of radius 1/2 sharing the geometrical centers. Randomly dropping points uniformly on the square, and counting the number of points landing on the disk, we can measure $\pi$ (or rather $\pi/4$ directly). How many points do we need to obtain 3 digits below the decimal point of $\pi/4$ with probability more than 99%?

**1.7** [Law of large numbers does not hold, if the distribution is too broad (if fluctuations are too wild)]
The Cauchy distribution that is defined by the following density distribution function

$$p(x) = \frac{1}{\pi}\frac{a}{x^2 + a^2} \tag{1.P.4}$$

does not satisfy $E(|X|) < +\infty$ (needless to say, the variance is infinite). Actually, the density distribution of

$$E_n = \frac{X_1 + \cdots + X_n}{n} \tag{1.P.5}$$

has exactly the same distribution function as $X_1$, if $\{X_j\}$ all obey the same Cauchy distribution and are statistically independent. Let us demonstrate this.
(1) What is the characteristic function of the Cauchy distribution? You can look up the result, but even in that case you must explain why the result is correct.
(2) Show what we wish to demonstrate.

**1.8** [St. Petersburg Paradox by Daniel Bernoulli]
Let $\{X_i\}$ be iid with

$$P(X_1 = 2^n) = 2^{-n} \tag{1.P.6}$$

---

[138] 4.16 of J. P. Romano and A. F. Siegel, *Counterexamples in Probability and Statistics*(Wadsworth & Brooks/Cole 1986).

for all positive integers $n$.
(1) Show that $E(X_1) = \infty$.
Thus, it seems that if $X_1$ is the gambler's gain, the reader can participate in this gambling game with any entry price and still can expect a positive gain. However, any sensible person would pay \$1000 as a fair price for playing. Why? This is the 'paradox.'
(2) Needless to say, the law of large numbers does not hold for $E_n$. This implies that empirically obtainable expectation and theoretical one should have some discrepancy. Indeed, it can be proved (the reader need not show this; not very easy) that for any positive $\varepsilon$

$$P\left(|E_n/\log_2 n - 1| > \varepsilon\right) \to 0 \tag{1.P.7}$$

in the $n \to \infty$ limit. Recall that $E_n$ is the expected payoff. Explain why the reader does not wish to pay \$1000. (Or for this to be a fair price how many times does the reader have to play?)

**1.9** [Bertrand's paradox]
Explain 'Bertrand's paradox' in about 10 lines (without using outrageous fonts). What lesson can you learn? [You can easily find a detailed account in the web.]

**1.10** [System with dissipation]
There is a classical particle system described by the canonical coordinates $\{q, p\}$ ($q$ and $p$ are collective notations for position and momentum coordinates of all the particles in the system). In terms of the Poisson bracket and the system Hamiltonian $H$ the equation of motion may be written as

$$\frac{dq}{dt} = [q, H], \quad \frac{dp}{dt} = [p, H] - \alpha p, \tag{1.P.8}$$

where $\alpha$ is a small positive constant. That is, the system is not perfectly classical mechanical, but contains a certain effect of dissipation.[139]
(1) Demonstrate that the Liouville's theorem is violated for this system.
(2) Demonstrate that the system energy decreases. (Assume that $H = K + U$ as usual and $K$ is quadratic in $p$.)
(3) However, show that if $H$ is time dependent, then the system energy may be maintained.

**1.11** [Wrong logic in a popular textbook]
The following can be read in a textbook. Point out the error in the argument.

> In general, there is no logical "room" for adding extra assumptions, such as equal a priori probability. The evolution of an actual system is determined by the laws of mechanics (or quantum mechanics). If the results

---

[139] This model may look artificial, but similar models are used to study nonequilibrium systems.

of using any extra assumptions always agree with the logical consequence of the laws of mechanics, and it should be possible to show that fact. If they do not agree with the laws of mechanics, then the extra assumptions are wrong.

**1.12** [Another expression of microcanonical partition function]
Classically, the microcanonical distribution may be written as

$$\hat{w}(E) = \frac{1}{N!h^{3N}} \int \delta(H(q,p) - E)dqdp. \tag{1.P.9}$$

Show that this can be expressed as follows:

$$\hat{w}(E) = \frac{1}{N!h^{3N}|grad\, H|} d\sigma, \tag{1.P.10}$$

where $\sigma$ is the area element of the constant energy surface and $grad\, H$ is the $3N$-dimensional gradient vector of the system Hamiltonian $H$ with respect to the canonical variables $(q_i, p_i)$.

**1.13** [Equipartition of energy with the aid of microcanonical ensemble]**r**
Within the microcanonical ensemble formalism[140] for a classical fluid consisting of $N$ interacting but identical particles,
(1) Show that the temperature $T$ can be expressed in terms of the average of the reciprocal kinetic energy as ($N \gg 1$ may be assumed)

$$k_BT = \frac{2}{3N\langle K^{-1}\rangle}, \tag{1.P.11}$$

where $K$ is the total kinetic energy of the system.
Comment: We are NOT discussing ideal systems. The system can have any interaction among particles.
(2) In the thermodynamic limit show that this is equivalent to

$$k_BT = \frac{2}{3}\langle\kappa\rangle, \tag{1.P.12}$$

where $\kappa$ is the average kinetic energy par particle. This is essentially the equipartition of kinetic energy. [Hint. the reader may use intuitively the weak law of large numbers.]

**1.14** [Generalized homogeneous function]

---

[140] Use $\Omega(E)$, the phase volume of the states with the total energy not exceeding $E$.

As we will learn much later, various thermodynamic quantities diverge at the second-order phase transition point (critical point). The free energy density $f$ as a function of temperature $\tau \propto T - T_c$ and pressure $p$ behaves as[141]

$$f(\lambda^{y_1}\tau, \lambda^{y_2}p) = \lambda^d f(\tau, p), \tag{1.P.13}$$

where $\lambda$ ($> 0$) is a scaling factor, $d$ is the spatial dimensionality, and $y_1$ and $y_2$ are real constants (related to the so-called critical exponents). That is, $f$ is a generalized homogeneous function.[142]

Formulate the counterpart of Euler's theorem and demonstrate it. You may freely use the method of characteristics.

**1.15**. [Mixing entropy and Gibbs' paradox]$_{\mathbf{r}}$
We have two ideal gases with the same volume $V$, pressure $P$, and temperature $T$. These two gases consist of different chemical species. Assume that the whole system is thermally isolated during the following processes.
(1) Two boxes containing the above gases are connected. That is, now the total volume of the mixture is $2V$. Find the entropy change due to this procedure of joining two boxes.
(2) Find the entropy change if two gases are mixed into a single volume of $V$.
(3) How can you actually measure the entropy change in (1) experimentally?

**1.16** [To check that Boltzmann's formula does not contradict thermodynamics]$_{\mathbf{r}}$
Let us check that Boltzmann's principle (within classical physics) is indeed consistent with thermodynamics: that is, if $S = k_B \log \hat{w}(E, V)$,

$$dS = \frac{dE + PdV}{T}, \tag{1.P.14}$$

where $\hat{w}(E, V)$ is the phase volume of microstates satisfying that the energy is in $(E - \delta E, E]$ and the volume is in $(V - \delta V, V]$. Here, we clearly know what $E$ and $V$ are in both mechanics and in thermodynamics. The pressure $P$ can be computed mechanically, and $T$ is related to the average kinetic energy $K$ of the system.

Using the Boltzmann formula

$$dS = k_B \frac{1}{\hat{w}} \frac{\partial \hat{w}}{\partial E} dE + k_B \frac{1}{\hat{w}} \frac{\partial \hat{w}}{\partial V} dV. \tag{1.P.15}$$

Therefore, if we can compute partial derivatives in the above and identify their meanings we should accomplish what we desire. This is actually what Boltzmann did in

---

[141] Precisely speaking, this is the singular part of the free energy as we will learn later. Peculiar phenomena near the critical point are governed by this part of the free energy.

[142] B. Widom realized from the empirical data that if $f$ is a generalized homogeneous function, then critical peculiar phenomena can be explained in a unified fashion.

1864. The demonstration is not very trivial, so we accept the following relation

$$k_B \frac{1}{\hat{w}} \frac{\partial \hat{w}(E,V)}{\partial V} \to \frac{P}{T} \tag{1.P.16}$$

(in the thermodynamic limit) and consider only the energy derivative. We can write

$$\hat{w}(E,V) = \int_{[E]} d\boldsymbol{q}d\boldsymbol{p} - \int_{[E-\delta E]} d\boldsymbol{q}d\boldsymbol{p} = \delta E \frac{\partial}{\partial E} \int_{[E]} d\boldsymbol{q}d\boldsymbol{p}, \tag{1.P.17}$$

where $[E]$ denotes the phase volume with energy not larger than $E$. We assume that the gas is confined in the volume $V$. Let $E = K(\boldsymbol{p}) + U(\boldsymbol{q})$, where $K$ is the total kinetic energy, and $U$ the total intermolecular potential energy. The phase integration may be written as

$$\int_{[E]} d\boldsymbol{q}d\boldsymbol{p} = \int d\boldsymbol{q} \int_{K(\boldsymbol{p}) \le E - U(\boldsymbol{q})} d\boldsymbol{p}. \tag{1.P.18}$$

Thus, the integration with respect to $\boldsymbol{p}$ is the calculation of the volume of the $3N$-sphere of radius $\sqrt{2m(E-U(\boldsymbol{q}))}$.
(1) Show that

$$\frac{\partial}{\partial E} \int_{[E]} d\boldsymbol{q}d\boldsymbol{p} = \int d\boldsymbol{q} \frac{S_{3N}}{3N} 2m \frac{3N}{2} [2m(E-U(\boldsymbol{q}))]^{3N/2-1}, \tag{1.P.19}$$

where $S_{3N}$ is the surface area of the $(3N-1)$-dimensional unit sphere.
(2) Using this formula, we can differentiate the integrand with $E$. Obtain

$$\frac{1}{\hat{w}(E,V)} \frac{\partial \hat{w}(E,V)}{\partial E} = k_B \left( \frac{3N}{2} - 1 \right) \left\langle \frac{1}{K(\boldsymbol{p})} \right\rangle. \tag{1.P.20}$$

(3) We know from the kinetic theory that the average kinetic energy of a point particle is proportional to $T$ (precisely speaking, the average of $\boldsymbol{p}^2/2m = 3k_BT/2$). Assuming that all the kinetic energies of the particles are statistically independent,[143] demonstrate that the formula in (2) is indeed equal to $1/T$.

**1.17** [Superrelativistic ideal gas]
Consider a super-relativistic gas consisting of particles whose energy $\varepsilon = c|\boldsymbol{p}|$, where $c$ is the speed of light, and $\boldsymbol{p}$ is the particle translational momentum.

---

[143] This is not really a trivial statement; we need that the system is 'normal.' That is, the intermolecular interaction range must be very short, and the interactions are sufficiently repulsive in the very short range.

(1) We have learned that the equation of state and the constant volume specific heat $C_V$ of an ideal gas may be obtained almost dimensional analytically, if we accept the basic postulate of statistical mechanics. Following this logic, find the pressure and $C_V$.
(2) Calculate the entropy to determine the constant corresponding to '$c$' (the constant in the entropy formula).

**1.17'** [Generalization of 1.17]
There is an ideal gas consisting of $N$ particles whose kinetic energy is given by $\varepsilon = c|\boldsymbol{p}|^\alpha$, where $\boldsymbol{p}$ is the momentum, and $\alpha$ and $c$ are positive constants ($\alpha = 1$ is the superrelativistic case). Answer the following questions with dimensional analysis (+ perhaps a little of thermodynamic principles, if needed). [Explicit calculations will cost you significant points.]
(1) What is the equation of state?
(2) What is the pressure as a function of internal energy?

**1.XX** [Frenkel defect]
The atoms can be dislocated into interstitial sites upon heating a crystal. They are the *Frenkel defects*. Suppose each atom can occupy either a regular lattice site or an interstitial site. Let us assume that for each lattice site there are $b$ interstitial sites. Assume that the crystal is made of $N$ atoms. Here, we assume that the state of each atom does not depend on its environment.
(1) What is the entropy of the state with $n$ dislocated atoms? (You need not try to simplify the result.)
(2) Suppose it costs an energy $\varepsilon$ for an atom to move into an interstitial site from a regular site. What happens to $n$, if we double the number of available interstitial sites? You may assume $\varepsilon/k_BT$ is large enough.
(3) Instead of (2) what happens if $\varepsilon = 0$?

**1.18** [Application of the Sackur-Tetrode equation[144]]
The following data are all under 1 atm.
The melting point of mercury is 234.2 K and the heat of fusion is 2.33 kJ/mol. The absolute entropy of solid mercury just at the melting temperature is 59.9J/K·mol. The entropy increase of liquid between the melting point and the boiling point is 26.2 J/K·mol. The boiling point is 630 K and the evaporation heat is 59.3 kJ/mol.
(1) Calculate the absolute entropy of mercury gas just above the boiling point.
(2) Assuming that mercury vapor is a monatomic ideal gas, estimate Planck's con-

---

[144] The best elementary exposition is found in F. Mandl, *Statistical Physics* (Wiley, 1971) Section 7.6.

stant. The reader may use the value of $k_B$.

**1.19** [Negative temperature]
Let us consider the two state spin system containing 1 mole of spins as discussed in the text. Assume that under the magnetic field $h$, the energy gap between the up and down spin states is 600 K per spin. Suppose the initial temperature of the magnet is $-500K$.
(1) What is the temperature of this system measured with an ideal gas thermometer containing $10^{-6}$ moles of monatomic gas particles?
(2) If, instead, the original negative temperature system is thermally equilibrated with a 1 mole of ideal monatomic gas that is initially 200K, what is the equilibrium temperature?

# Chapter 2

# Canonical Distribution

## 2.1 Isothermal systems

Isolated (or thermally isolated) systems are not very convenient practically. Often we study systems whose temperature is maintained (thermostatted). This is realized by thermal contact between the system and a heat bath; a heat bath is a sufficiently large system that can equilibrate fast, so its temperature does not change by its interaction with the system of our interest.

Do we clearly know all the words in the above paragraph? What is 'thermal contact'? It is a contact that allows heat exchange. Then, what is heat? We will learn that our incomplete understanding of heat is a source of big trouble. To characterize heat microscopically could be a major (perhaps the greatest) fundamental problem in statistical physics.

What is heat? It is defined only macroscopically (at least in the current state of physics[1]). Equilibrium states of a macroscopic system can be specified uniquely by internal energy and work coordinates (see [T5] in the preceding Chapter). For a pair of distinct such states, it is possible to change one of them to the other solely by supplying work from outside under the condition that the system can be influenced

---

[1] However, if a system may be described in terms of Langevin equations, there is a systematic way to define heat. See K. Sekimoto, *Stochastic Energetics* (Springer, 2010).

only by work (adiabatic condition) (this is a part of the second law; the process is generally not quasistatic). This allows us to measure the internal energy difference $\Delta E$ between these two states. Next, let us change the initial state to the final one without adiabatic condition. In this case, even if the process is not quasistatic, still the work $W$ added to the system (or done by the system; it is only a matter of the sign) can be quantified with the aid of mechanics. The discrepancy $Q = \Delta - W$ between $\Delta E$ and $W$ is called *heat*.[2] This is an operational definition, so a very respectable definition in physics. The wall that allows $Q \neq 0$ to happen is called a *diathermal wall* and the contact between two systems through such a wall without any exchange of quantities described by work coordinates is the *thermal contact*. A heat bath is a constant temperature system contacting with a system under study through a diathermal wall.[3] A wall that does not allow heat exchange but allows exchange of work through it is called an *adiabatic wall*. An adiabatic process is a process that may occur to a system enclosed by an adiabatic wall.[4]

If we take a big uniform isolated macroscopic object and surround its small (but still macroscopic) part with a diathermal wall that does not allow exchange of work coordinates, we can imagine the whole system as a system of our interest (System I) + its surrounding heat bath (System II). Even if we must discuss a macrostate in a thermally isolated system that requires external fields, we need only obvious modifications, so we discuss the case of isolated systems as in the standard textbooks.

---

[2] If the process adding work $W$ is not a quasistatic reversible process, the external work added to the system can be dissipated in the system to generate heat, so $W$ cannot be obtained in terms of the work coordinate changes. If the work is added reversibly and quasistatically, it can be obtained by integrating $\sum x_i dX_i$ along the process.

[3] In practice, we may use a device that can control the temperature with the aid of a feedback mechanism while monitoring the system temperature.

[4] ⟪**Adiabatic condition**⟫ When we discuss an adiabatic system, the reader may at once imagine a system thermally isolated in a Dewar jar. It may be a good picture of a thermally isolated system, but in thermodynamics an adiabatic process is a process for which no net heat exchange exists between the system and any heat bath. For example, if a system is attached to a heat bath and if there is no net exchange of heat between the system and the heat bath, it is a respectable adiabatic process. Although an isolated system may be described purely in terms of mechanics, there is no way to describe a general adiabatic process in terms of pure mechanics.

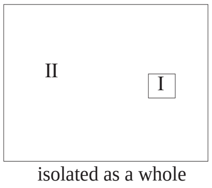


Fig. 2.1.1
Heat bath (II) and the system (I).

rLet $E_X$ ($X =$ I or II) be the internal energy of System $X$, and $E_{\rm I+II}$ be the total internal energy of the thermally contacting I and II.[5] Since the systems are assumed to be macroscopic, we may ignore the energy change due to the contact of these two systems:

$$E_{\rm I+II} = E_{\rm I} + E_{\rm II}. \tag{2.1.1}$$

The ground state energy of the system is set to be zero. Let $w_{\rm II}(E)$ be the number of microstates for System II with energy in $(E - \delta E, E]$.[6] Since the interaction between I and II is weak, we may assume that these systems are statistically independent, so the number of microstates allowed to the total system is

$$w_{\rm I+II}(E) = \sum_{0 \le E_{\rm I} \le E} g(E_{\rm I}) w_{\rm II}(E - E_{\rm I}). \tag{2.1.2}$$

Here, the sum on the right-hand side is taken over all microstates of I (all the eigenstates of the Hamiltonian of isolated I) whose energy $E_{\rm I}$ belongs to $[0, E]$. $g(E_{\rm I})$ denotes the multiplicity of such microstates. Therefore, assuming that the principle of equal probability holds for the total system, the probability $f(E_{\rm I})$ that System I has energy $E_{\rm I}$ is given by

$$f(E_{\rm I}) = \frac{1}{w_{\rm I+II}(E)} g(E_{\rm I}) w_{\rm II}(E - E_{\rm I}). \tag{2.1.3}$$

Applying Boltzmann's principle, we may rewrite (2.1.3) in terms of the entropy of the heat bath $S_{\rm II}(E_{\rm II})$ as

$$f(E_{\rm I}) \propto g(E_{\rm I}) \exp\left\{ \frac{1}{k_B} S_{\rm II}(E - E_{\rm I}) \right\}. \tag{2.1.4}$$

Let us consider the following Taylor expansion:

$$S_{\rm II}(E - E_{\rm I}) = S_{\rm II}(E) - E_{\rm I} \frac{\partial S_{\rm II}}{\partial E_{\rm II}} + \frac{1}{2} E_{\rm I}^2 \frac{\partial^2 S_{\rm II}}{\partial E_{\rm II}^2} + \cdots. \tag{2.1.5}$$

---

[5] For general thermally isolated systems, the following discussion would be slightly more involved, because the energy of the system and its internal energy may not be identified.

[6] We have already discussed that $\delta E$ may be chosen appropriately, but in practice, it is extensive, i.e., it is a very small but a finite fraction of $E$.

$S_{\rm II}$ and $E_{\rm II}$ are both extensive, so $\partial S_{\rm II}/\partial E_{\rm II} = O[N_{\rm II}^0]$ (i.e., intensive), where $N_{\rm II}$ is the number of particles in System II. Furthermore, $\partial^2 S_{\rm II}/\partial E_{\rm II}^2$ must be the homogeneous function of degree $-1$ of the number of particles, so we conclude that $\partial^2 S_{\rm II}/\partial E_{\rm II}^2 = O[N_{\rm II}^{-1}]$. This implies that the second term on the RHS of (2.1.5) is $O[N_{\rm I}]$ and the third term is $O[N_{\rm I}^2/N_{\rm II}]$. Therefore, relative to the second term, we may ignore the third term. (2.1.4) may be written as

$$f(E_{\rm I}) = \frac{1}{Z} g(E_{\rm I}) e^{-E_{\rm I}/k_BT}, \tag{2.1.6}$$

where we have used the definition of the absolute temperature $T$. $Z$ is a normalization constant called the *canonical partition function*:

$$Z = \sum_E g(E) e^{-\beta E}. \tag{2.1.7}$$

Here, the summation is over all the distinct eigenvalues of the Hamiltonian $H$ of System I, and $\beta$ is $1/k_BT$ as usual. Since $g(E)$ is the multiplicity of the eigenvalue, we may write

$$Z = Tr\, e^{-\beta H}. \tag{2.1.8}$$

The obtained probability distribution (2.1.6) of energy of System I is called the *canonical distribution*. The probability of the microstates with the same energy must be the same according to the principle of equal probability, so we may write the corresponding density operator (*canonical density operator*) as

$$\rho = \frac{1}{Z} e^{-\beta H}. \tag{2.1.9}$$

Classically, if the Hamiltonian of the $N$ particle system is $H$, the 'multiplicity' $g(E)$ appearing in the above should be interpreted as the phase volume of a thin energy shell $(E - \delta E, E]$ of width $\delta E$, so we should write it as

$$\tilde{w}(E) = \frac{1}{N! h^{3N}} \int_{E-\delta E \le H < E} d\Gamma_N, \tag{2.1.10}$$

where $d\Gamma_N$ is the volume element of the phase space. Therefore, the canonical partition function may be written as

$$Z = \frac{1}{N! h^{3N}} \sum_E \int_{E-\delta E \le H < E} d\Gamma_N\, e^{-E/k_BT} = \frac{1}{N! h^{3N}} \int d\Gamma_N\, e^{-H/k_BT}. \tag{2.1.11}$$

Here, the summation over $E$ actually means to collect all the energy shells of thickness $\delta E$. Because the probabilities of the microstates with the same energy are the same (the principle of equal probability), the probability density for the microstate $\gamma$ may be written as

$$P(\gamma) = \frac{1}{N!h^{3N}Z}e^{-\beta H(\gamma)}, \tag{2.1.12}$$

where $H(\gamma)$ is the energy of the microstate $\gamma$. This is the (classical) canonical distribution of the microstates.

Thus, the canonical partition function $Z$ may be written generally as**r**

$$Z = \sum_E w(E)e^{-\beta E}, \tag{2.1.13}$$

where $w(E)$ is the number of states in the energy shell $(E - \delta E, E]$ (if classical, it is replaced with the corresponding $\tilde{w}$ (2.1.10)). $w(E)$ is the microcanonical partition function introduced in Chapter 1, so the canonical partition function is**r** essentially the Laplace transform of the corresponding microcanonical partition function.[7]

Since $w(E)$ is usually a rapidly increasing function, $w(E)e^{-\beta E}$ has a peak (this also applies to $\tilde{w}(E)$). $E$ is extensive and $\log W(E)$ is also extensive, so the peak must be a very sharp peak.[8] Therefore,

$$Z \simeq w(E^*)e^{-\beta E^*}, \tag{2.1.14}$$

where $E^*$ is the energy that maximizes $w(E)e^{-\beta E}$. Or, we may write[9] (2.1.14) as

$$Z \simeq \exp\left[\frac{1}{k_B}\sup_E\left(S(E) - E/T\right)\right]. \tag{2.1.15}$$

Therefore,**r**

$$\inf_E(E - TS) = A = -k_BT\log Z. \tag{2.1.16}$$

Thus, $Z$ directly gives the Helmholtz free energy $A$. This is the algorithm mentioned at the beginning of this book.

---

[7] Here, following the tradition, no work coordinates are explicitly written, but all the work coordinates required to single out the relevant thermodynamic state must be fixed (with the usual reasonable leeway).

[8] This is the chief idea behind the WKB or the saddle point method (or Varadhan's theorem in large deviation theory).

[9] ⟪**Sup and inf**⟫ sup implies '*supremum.*' $\sup A$ denotes the smallest number that is larger than any number in $A$. The reader may intuitively understand that it is the maximum in $A$ (the reason why such a non-intuitive concept is needed is that $A$ may not be a closed set). Analogously, inf is '*infimum*' and the reader may interpret $\inf A$ roughly as the minimum of $A$.

This concludes the construction of the statistical framework for isothermal systems. We started from the principles for isolated systems,[10] but we may start directly with isothermal systems.

To construct statistical mechanics, we do not need any statistical principle, but have only to specify how to compute thermodynamic observables in terms of mechanics (the correspondence rules). To this end, as before, we need only the correspondence rule for thermodynamic coordinates and that for a thermal quantity. For thermodynamic coordinates, the correspondence rules are the same as in the microcanonical case. For the thermal correspondence we have only to postulate

$$A = -k_B T \log Z \tag{2.1.17}$$

instead of Boltzmann's principle, where for $Z$ we adopt the form as given above. In place of postulating this correspondence rule, we should be able to postulate a statistical principle: the probability distribution for the microstate $\gamma$ is given by (2.1.12) classically, and the density operator is given by (2.1.9) quantum-mechanically. This was the approach chosen by Einstein as we will see soon below.

We have seen in Section 1.1 that thermodynamics requires only instantaneous observations of macroobservables, so we need only a representative microstate. The reproducibility of thermodynamic observables is likely due to statistical independence of remote parts of a macroscopic system. Therefore, if a system is sufficiently large, then a single energy eigenstate[11] must be described locally as a canonical distribution.[12]

Let us look at Einstein's approach$_{\mathbf{r}}$ to statistical mechanics. The correspondence rule for the thermodynamic coordinates is the same as usual. From this interpretation

---

[10] To be precise, the formalism obtained here can always be used and gives equivalent results to those obtained by the microcanonical ensemble method (see Sect. 2.6), but the derivation cannot always be justified. Therefore, the argument given here should be regarded as a heuristic argument to motivate the canonical formalism. The argument we used can be found in Einstein's paper of 1903, which is almost his first paper; It can also be found in Gibbs' famous book.

[11] We wish to say 'a single quantum-mechanically pure state' instead of 'a single energy eigenstate,' but as noted already there are many pathological pure states, so to be safe we confine ourselves to 'energy eigenstates.'

[12] This argument has been justified: e.g., S. Goldstein, J. L. Lebowitz, R. Tumulka and V. N. Zangih, "Canonical typicality," Phys. Rev. Lett. **96**, 050405 (2006). However, needless to say, microcanonical distribution is assumed. This assumption cannot be removed as we have discussed at length. The physical essence of statistical mechanics lies in that 'almost any' single energy eigenstate is interpretable as a microcanonical ensemble.

of $dE$ and $d'W$ in terms of mechanics is given.[13] Einstein chose $P = (1/Z)e^{-\beta H}$ for his statistical principle. We assume that the system Hamiltonian depends on parameters $\lambda_i$**r** that may be controlled externally. Let us write the change of the Hamiltonian due to the change of these parameters as $\delta H$, and the rest as $d'H$:[14]

$$dH = \delta H + d'H. \tag{2.1.18}$$

Since the control parameters can be changed adiabatically, Einstein concludes that the first term in (2.1.18) (averaged over the canonical distribution) is the work. Therefore, inevitably, heat corresponds to

$$d'Q \equiv \langle d'H \rangle. \tag{2.1.19}$$

This is his translation rule (but not independent of the translation rule for thermodynamic coordinates).

Let $P = e^{-\log Z - \beta H}$ be the canonical distribution. Let us change the temperature (i.e., $\beta$) and the mechanical parameters $\{\lambda_i\}$:

$$0 = d\int P dy = \int dy\, [-d\log Z - H d\beta - \beta\delta H]\, P \tag{2.1.20}$$

Here $d$ is the total derivative with respect to the variables $\beta$ and $\{\lambda_i\}$. This reads

$$0 = -d\log Z - E d\beta - \beta d'W. \tag{2.1.21}$$

Using the microscopic expression of heat (2.1.19), we may rewrite this as

$$0 = -d\log Z - E d\beta - \beta dE + \beta d'Q = -d\log Z - d(\beta E) + \beta d'Q \tag{2.1.22}$$

That is,

$$\beta dQ = d\log Z + d(\beta E). \tag{2.1.23}$$

According to thermodynamics, this reads

$$dS = \frac{1}{T} dQ = d\left(\frac{E + k_B T\log Z}{T}\right) \tag{2.1.24}$$

---

[13] If we use the ordinary correspondence $E = \langle H \rangle$, the Gibbs-Helmholtz relation (a purely thermodynamic relation) gives $-\partial \log Z/\partial\beta = \partial(\beta A)/\partial\beta$. This is, however, a relation between partial derivatives, so we cannot conclude $A = -k_B T\log Z$ from this.

[14] Einstein wrote explicitly as

$$\delta H = \sum_i \left(\frac{\partial H}{\partial \lambda_i}\right)_{q,p} \delta\lambda_i,$$

but such an expression is not needed.

or $-k_B d\log Z = d(A/T)$. Integrating this and ignoring the arbitrary integration constant,[15] Einstein arrived at

$$A = -k_B T \log Z. \tag{2.1.25}$$

Notice that the above argument does not depend on any explicit formula of heat; what is needed is in (2.1.22) $\delta H$ is the systematic change due to work.

r Since $-k_B d\log Z = d(A/T)$ is an equation obtained under the constant particle number $N$ condition, (2.1.25) actually must include an arbitrary function $f(N)$ of $N$ as $A = -k_B T \log[f(N)Z]$. As stated when we discussed the Gibbs paradox, there is no way to determine this for a system without change of $N$ (nor there is any necessity to do so). What we saw before was that the factor $f(N)$ is determined if we demand the consistency with the fourth law of thermodynamics, when $N$ may be changed.

**Derivation of the canonical distribution, the general case**

As stated before, if we consider only isolated system not subjected to external influences, we cannot have a statistical theory applicable to arbitrary thermodynamic states. We must consider thermally isolated systems as well. We realized the states we must count must be modified from the isolated cases. This modification propagates to the derivation of canonical distribution. Although here the case of thermal isolation in general is discussed, the outcome is quite parallel to the isolated case; simply replace the system Hamiltonian $H$ with the Hamiltonian $H + H_i$, where $H_i$ is the interaction Hamiltonian of the system with the external fields.

The stage of the derivation is the same as Fig. 2.1.1. The thermally isolated system is assumed to be described by the Hamiltonian $H_{\mathrm{I+II}} = H + H_{\mathrm{II}}$ (here the suffix I is dropped for System I). Let the Hamiltonian of the system M needed to impose the field(s) be $H_M$, and the interaction Hamiltonian between System M and System I (resp., System II) be, $H_i$ (resp., $H_{i,\mathrm{II}}$). The total Hamiltonian we must consider is the sum of all these. However, System M is assumed to be a classical macroscopic system, we need not worry about it to develop statistical mechanics, so the total Hamiltonian we must consider is

$$H + H_i + H_{\mathrm{II}} + H_{i,\mathrm{II}}. \tag{2.1.26}$$

We have already seen that Boltzmann's principle holds, if we replace internal energy with an appropriate thermodynamic potential adapted to the thermally isolated system. This implies that the argument used for isolated systems directly translates by replacing internal energy $E$ with an appropriately generalized enthalpy $J$ (such that $dJ = TdS - Xdx + \cdots$), because $J$ is the expectation value of $H_T = H + H_i$ under the principle of equal probability of its eigenstates.

Therefore, we have only to rewrite (2.1.2) as

$$w_{\mathrm{I+II}}(J) = \sum_{0\le J_{\mathrm{I}}\le J} g(J_{\mathrm{I}})w_{\mathrm{II}}(J - J_{\mathrm{I}}), \tag{2.1.27}$$

---

[15] This is allowed, because we have only to shift the origin of the energy.

where the sum is over the microstates of System I which are the eigenstates of $H_T = H + H_i$ whose eigenvalues $J_{\mathrm{I}}$ are in $[0, J]$, and $g(J_{\mathrm{I}})$ is the multiplicity of the eigenvalues. The subsequent argument with the aid of Taylor expansion is quite parallel to the case of isolated systems, and we obtain the formula corresponding to (2.1.5):

$$f(J) = \frac{1}{\tilde{Z}} g(J) e^{-J/k_B T}. \tag{2.1.28}$$

Do not forget that $g(J)$ is the number of eigenstates belonging to the eigenvalue $J$ of $H_T$ (System I interacting with M). The partition function $\hat{Z}$ may be written as

$$\hat{Z} = Tr\, e^{-H_T/k_B T}. \tag{2.1.29}$$

The obtained ensemble is under constant temperature and some external fields, so the natural thermodynamic potential is the generalized Gibbs free energy $\hat{G} = J - TS$:

$$\hat{G} = -k_B T \log \tilde{Z}. \tag{2.1.30}$$

For example, for a magnet under a constant field, $\hat{G} = A - BM$, and the Gibbs relation is $d\hat{G} = -SdT - MdB$.

### †Legendre transformation rq

The Helmholtz free energy $A$ is defined by $A = \inf_E\{E - TS\}$. This may be understood as $-A = \sup_S\{TS - E\}$, because $E$ is monotonically related to $S$. Such a transformation called the Legendre transformation appears again and again in thermodynamics.

The true essence of the *Legendre transformation* is: a convex curve can be reconstructed from a set of its tangent lines, where a tangent line of a convex curve is a line sharing at least one point with the curve, and all the points on the curve are on one side of the line or on it (i.e., none on the other side).[16]

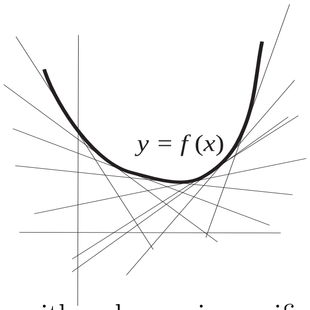


Fig. 2.1.2
The totality of tangent lines can recover a convex function. The picture can easily be generalized to higher dimensional spaces with the aid of tangent hyperplanes.

A line with a slope $\alpha$ is specified by its $y$-section $-f^*(\alpha)$: $y = \alpha x - f^*(\alpha)$. If this line is tangent to $f$, $f^*(\alpha)$ is given by the *Legendre transformation* of $f$ (Fig. 2.1.3):

$$f^*(\alpha) = \sup_x [\alpha x - f(x)]. \tag{2.1.31}$$

---

[16] A line is also understood as (a limiting case of) a convex curve. Accordingly, we understand that a function that is everywhere infinite except at one point is also a limiting case of a convex curve.

This is the mathematically standard definition of the Legendre transformation $f \to f^*$.

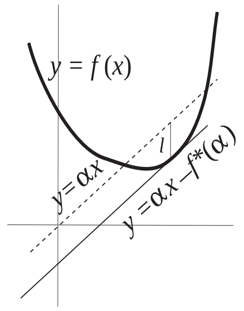


Fig. 2.1.3 $l$ is the maximum gap between the dotted line $y = \alpha x$ and the convex curve $y = f(x)$ (we pay attention to its sign; maximum of $\alpha x - f(x)$). Therefore, if we choose

$$f^*(\alpha) = \sup_x[\alpha x - f(x)],$$

then $y = \alpha x - f^*(\alpha)$ is the tangent line in the figure. This gives a geometrical meaning of the Legendre transformation $f \to f^*$.

If $f$ is convex, then $f^*$ is convex, and $f^{**} = f$. That is, the inverse Legendre transformation may be given by a symmetric procedure $f(x) = \sup_\alpha[\alpha x - f^*(\alpha)]$. This can be illustrated by Fig. 2.1.4. This graphic demonstration uses the fact that any convex function is a primitive function of an increasing function $g$: $f(x) = \int^x g(x')dx'$.

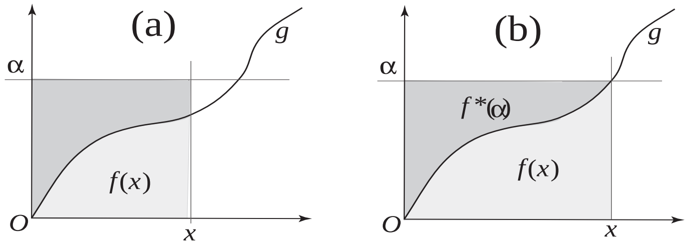


Fig. 2.1.4 Illustration of the relation between $f$ and $f^*$ in 1D.

In (a) of Fig. 2.1.4 the pale gray area is $f(x)$. Legendre transformation maximizes the signed area $\alpha x - f(x)$, the dark gray area, by changing $x$, that is, the (signed) area bounded by the $\alpha$-axis, the horizontal line through $\alpha$, the vertical line through $x$, and the graph of $g(x)$. When $\alpha = g(x)$, this dark gray area becomes maximum. This is realized in (b): $f^*(\alpha) + f(x) = \alpha x$ (this equality is called *Fenchel's equality*; obviously for arbitrary $x$ and $\alpha$ $f^*(\alpha) + f(x) \geq \alpha x$ (*Young's inequality*)). From these illustrations it should be obvious that the relation between $f$ and $f^*$ is perfectly symmetric, so $f^*$ is convex, and $f(x) = \sup_\alpha[\alpha x - f^*(\alpha)]$. Although such an illustration is impossible for higher dimensional case, still the relation between $f$ and $f^*$ is correct, so are Fenchel's equality and Young's inequality.[17]

As we will see $E$ is a convex function of $S$ (specific heat must be positive), so $-A = \sup_S\{ST - E\}$ is a convex function of $T$. That is, $A$ be must convex upward

---

[17] ⟪**Convex analysis bible**⟫ The classic text book of *convex analysis* is R. T. Rockafellar, *Convex Analysis* (Princeton UP, 1970; reprinted as a volume in the series Princeton Landmarks of Mathematics in 1997).

as a function of $T$.

We discussed only the 1-dimensional situation above, but Legendre transformation may be defined in multidimensional space as well: $f^*(\boldsymbol{\alpha}) = \sup_{\boldsymbol{x}}[\boldsymbol{\alpha}\cdot\boldsymbol{x} - f(\boldsymbol{x})]$. Many theorems are common to the 1-dimensional case. In particular, if $f$ is convex, so is $f^*$. However, if Legendre transformation is not applied to the whole space spanned by $\boldsymbol{x}$, convexity is not necessarily preserved. Suppose we define $\sup_{x_1}[\alpha_1 x_1 - f(\boldsymbol{x})]$ with respect to the first component of $\boldsymbol{x}$. If $f$ is convex, then $g(\alpha_1, x_2, \cdots, x_n)$ is a convex function of $\alpha_1$ if $x_2, \cdots, x_n$ are fixed, but $g$ is not generally a convex function anymore as a multivariate function.

## 2.2 Simple applications of canonical distribution

As we will learn in Sect. 2.6, the microcanonical formalism and the canonical formalism are equivalent in the thermodynamic limit, so to compute thermodynamic quantities we may use a convenient formalism. Often, the canonical formalism is easier than the microcanonical formalism.

This point can be easily illustrated by the study of (classical) ideal gas. We can totally avoid the calculation of the volume of a high-dimensional sphere.

$$Z = \frac{1}{N!}\sum_{\{\boldsymbol{n}_i\}}\exp\left(-\beta\sum_i\frac{h^2\boldsymbol{n}_i^2}{8mL^2}\right) = \frac{1}{N!}\left[\sum_{\boldsymbol{n}}\exp\left(-\beta\frac{h^2\boldsymbol{n}^2}{8mL^2}\right)\right]^N = \frac{1}{N!}\left[\sum_{n=1}^{\infty}\exp\left(-\beta\frac{h^2n^2}{8mL^2}\right)\right]^{3N}. \tag{2.2.1}$$

The sum may be replaced by the integral:

$$\frac{1}{2}\int_{-\infty}^{\infty}\exp\left(-\beta\frac{h^2n^2}{8mL^2}\right)dn = \sqrt{\frac{2\pi mk_BTL^2}{h^2}} = \frac{\sqrt{2\pi}}{\lambda_T}L, \tag{2.2.2}$$

where $\lambda_T$ is the de Broglie thermal wave length. Therefore,**q**

$$Z = \frac{1}{N!}\left[\left(\frac{2\pi mk_BT}{h^2}\right)^{3/2}V\right]^N = \frac{1}{N!}\left[\frac{\sqrt{2\pi}}{\lambda_T}L\right]^{3N} = \left[\frac{en_QV}{N}\right]^N, \tag{2.2.3}$$

where we have used Stirling's formula and $n_Q = (\sqrt{2\pi}/\lambda_T)^3$.

We obtain

$$A = Nk_BT\log(n/n_Q) - Nk_BT. \tag{2.2.4}$$

This is of course consistent with $A = N\mu - PV$. $Z(T)$ may be considered as the generating function of $w(E)$ (or the Laplace transformation of the latter as noted already), so there is no wonder about the ease of the calculation. It is just the standard trick.

Let us look at the general structure of the canonical partition function $Z$ of a classical system consisting of $N$ interacting classical particles. The Hamiltonian of such a system has the following structure:

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + U(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N). \tag{2.2.5}$$

Here, $m_i$ is the mass of the $i$th particle, $\boldsymbol{p}_i$ its momentum vector, $\boldsymbol{r}_i$ its position vector, and $U$ the total potential energy of the system. The integral over the phase space appearing in the computation of the canonical partition function may be separated into the integral over the subspace spanned by the momentum vectors and that over the subspace spanned by the position vectors called the *configuration space*. Let $K$ be the total kinetic energy of the system. Then, we have

$$Z = \frac{1}{N!h^{3N}} \int d\Gamma_N \, e^{-\beta(K+U)} = \frac{V^N}{N!h^{3N}} \int d^{3N}\boldsymbol{p} \, e^{-\beta K} \frac{1}{V^N} \int d^{3N}\boldsymbol{r} \, e^{-\beta U}. \tag{2.2.6}$$

Therefore, we can write

$$Z = Z_{ideal} Q, \tag{2.2.7}$$

where

$$Z_{ideal} = \frac{V^N}{N!} \left( \frac{2\pi m k_B T}{h^2} \right)^{3N/2} \tag{2.2.8}$$

is the partition function for the classical ideal gas, and $Q$ is called the *configurational partition function*:

$$Q = \left\langle e^{-\beta U} \right\rangle_V . \tag{2.2.9}$$

Here, $V$ is the system volume and

$$\langle * \rangle_V = \frac{1}{V^N} \int_V \cdots \int_V d\boldsymbol{r}_1 \cdots d\boldsymbol{r}_N \; * \, . \tag{2.2.10}$$

That is, the configurational partition function is the spatial average of $e^{-\beta U}$. In classical statistical mechanics how to compute $Q$ is the problem.

Let us not dwell on very elementary examples; there are tons of problems in the standard textbooks and exercise books.[18] A few such problems can even be found

---

[18] The best exercise book is, R. Kubo, H. Ichimura, T. Usui and N. Hashitsume, *Statistical Mechanics* (North-Holland, 1990).

at the end of this chapter. These problems the reader need not be able to solve (quickly) when s/he becomes a real scientist; after all, the solutions are in the books and the reader's business will be to find new problems and to solve them. So, here let us discuss a few general conclusions we can easily obtain by the canonical formalism.

**Simple worked-out examples**

How to use the canonical formalism may be best reviewed through simple problems. Therefore, here a couple of elementary examples are solved in detail, and then uninteresting qual style problems will be listed.

Example 1 [Harmonic oscillators]
Compute the entropy of a collection of $N$ noninteracting 1D harmonic oscillators both classically and quantum-mechanically. What is the chief difference, if any, between these two entropies?

The classical partition function is (let us assume that $N$ oscillators are distinguishable)

$$Z_C = \left(\frac{1}{h}\int dpdq\, e^{-(1/2k_BT)[p^2/m+kq^2]}\right)^N = \left(\frac{1}{h^2}\frac{2\pi m}{\beta}\frac{2\pi}{\beta k}\right)^{N/2} = \left(\frac{k_BT}{\hbar\omega}\right)^N, \tag{2.2.11}$$

where $\omega = \sqrt{k/m}$. The quantum version is**q**

$$Z_Q = \left(\sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)}\right)^N = \left(\frac{1}{2\sinh\beta\hbar\omega/2}\right)^N. \tag{2.2.12}$$

Entropy is, classically (it is easier to perform the following calculation, with the aid of equipartition of energy (Sect. 2.3), than differentiation),

$$S_C = (E-A)/T = [Nk_BT + Nk_BT\log(k_BT/\hbar\omega)]/T = Nk_B\log T + \text{ const.} \tag{2.2.13}$$

Quantum mechanically (honest differentiation is the simplest), it is

$$S_Q = Nk_B\left(\frac{1}{2}\beta\hbar\omega\coth(\beta\hbar\omega/2) - \log(2\sinh(\beta\hbar\omega/2))\right). \tag{2.2.14}$$

The chief difference is in the $T \to 0$ limit. Classically, entropy goes to $-\infty$ (inconsistent with thermodynamics), but quantum mechanically it is zero (consistent with thermodynamics). □

Example 2. $N$ molecules of the same chemical species are fixed on a lattice. Each molecule has two spins of spin 1/2, which interact in the same molecule according to the following spin Hamiltonian:

$$h = -J\sigma_1\cdot\sigma_2 - \mu\boldsymbol{B}\cdot(\sigma_1+\sigma_2). \tag{2.2.15}$$

Here, $J$ is the coupling constant, $\boldsymbol{B}$ is the external magnetic field pointing in the $z$-direction, and $\mu$ is the spin-magnetic moment ratio. The spins are quantum spins of total spin angular momentum 1/2. Find the specific heat and the magnetization at $\boldsymbol{B} = 0$ (due to the spin degrees of freedom of the system).

The system Hamiltonian is the sum of the above spin Hamiltonians for individual molecules: $H = \sum_{i=1}^{N} h_i$, where $h_i$ is the spin Hamiltonian for the $i$th molecule. The canonical partition function of the system is[19]

$$Z = Tr_{\{\cdots \sigma_{1i}, \sigma_{2i} \cdots\}} e^{-\beta \sum h_i}, \tag{2.2.16}$$

where the trace $Tr_{\{\cdots \sigma_{1i}, \sigma_{2i} \cdots\}}$ implies the sum over all the spin states for all the molecules. Since all the states of individual molecules appear once and only once,

$$Z = \prod_{i=1}^{N} \left[ Tr_{\sigma_{1i}, \sigma_{2i}} e^{-\beta h_i} \right]. \tag{2.2.17}$$

Therefore, introducing the partition function for each molecule

$$z = Tr_{\sigma_{11}, \sigma_{21}} e^{-\beta h_1}, \tag{2.2.18}$$

(here, the first molecule is chosen as the representative, but since all the molecules are identical, the suffix '1' will be henceforth dropped), we can write $Z = z^N$. To calculate this, it is probably the easiest to obtain the eigenstates of the molecular spin Hamiltonian. Since

$$h = -J\sigma_1 \cdot \sigma_2 - \mu \boldsymbol{B} \cdot (\sigma_1 + \sigma_2) = -\frac{J}{2} \left[ (\sigma_1 + \sigma_2)^2 - \sigma_1^2 - \sigma_2^2 \right] - \mu \boldsymbol{B} \cdot (\sigma_1 + \sigma_2), \tag{2.2.19}$$

we may rewrite $h$ with the aid of the total spin operator $\sigma$ and its $z$-component $\sigma_z$ as

$$h = -\frac{J}{2}[\sigma^2 - 3/2] - \mu B \sigma_z. \tag{2.2.20}$$

Here, we have used $\sigma_1^2 = \sigma_2^2 = 3/4$. The total spin quantum number is 1 or 0, so the eigenvalue of $\sigma^2$ is 2 or 0. If the system under study is as simple as in this case, it is convenient to make a table of states. $E$ in the table denotes the eigenvalue of $h$, that is, the energy of the eigenstate:

| $\sigma$ | $\sigma_z$ | $E$ |
|---|---|---|
| 1 | 1 | $-J/4 - \mu B$ |
| 1 | 0 | $-J/4$ |
| 1 | $-1$ | $-J/4 + \mu B$ |
| 0 | 0 | $3J/4$ |

Therefore, we obtain

$$z = 2e^{\beta J/4} \cosh \beta \mu B + e^{\beta J/4} + e^{-3\beta J/4},$$

---

[19] Precisely speaking, since the system is under an external magnetic field, the following $Z$ is not a canonical partition function, but a generalized canonical partition function, and the corresponding thermodynamic potential is not simply $A$ but $A - BM$, but here, we follow the usual abuse of the notations and concepts.

so the free energy reads:

$$A = -Nk_BT\log z = -Nk_BT\log\left(2e^{\beta J/4}\cosh\beta\mu B + e^{\beta J/4} + e^{-3\beta J/4}\right).$$

The heat capacity may be obtained from entropy $C = T\partial S/\partial T$ or from internal energy $C = \partial E/\partial T$. When we use the canonical distribution, the latter is simpler. $E$ may be obtained with the aid of the Gibbs-Helmholtz relation

$$E = \left.\frac{\partial A/T}{\partial 1/T}\right|_V = -\frac{\partial\log Z}{\partial\beta}, \tag{2.2.21}$$

or we can directly write it down as the average of the energy.[20] Entropy should not be computed through differentiation $-\partial A/\partial T$, but should be computed as $S = (E-A)/T$.

$$E = -\frac{3NJ}{4}\frac{e^{\beta J/4}-e^{-3\beta J/4}}{3e^{\beta J/4}+e^{-3\beta J/4}}, \tag{2.2.22}$$

so when $\boldsymbol{B} = 0$,

$$C = \frac{3NJ^2}{k_BT^2}\frac{e^{-\beta J/2}}{(3e^{\beta J/4}+e^{-3\beta J/4})^2}. \tag{2.2.23}$$

Incidentally, we have

$$S = (E-A)/T = Nk_B\log\left(2e^{\beta J/4}\cosh\beta\mu B + e^{\beta J/4} + e^{-3\beta J/4}\right) - \frac{3NJ}{4T}\frac{e^{\beta J/4}-e^{-3\beta J/4}}{3e^{\beta J/4}+e^{-3\beta J/4}}. \tag{2.2.24}$$

Magnetization $\chi$ requires computing magnetization $M$. This may be obtained more conveniently directly with the aid of its definition rather than differentiating free energy:[21]

$$M = N\mu\frac{2\sinh\beta\mu B}{2\cosh\beta\mu B + 1 + e^{-\beta\mu J}}. \tag{2.2.25}$$

From this for $\boldsymbol{B} = 0$

$$\chi = \frac{\partial M}{\partial B} = N\mu^2\beta\frac{2}{3+e^{-\beta\mu J}}. \tag{2.2.26}$$

□

Example 3. A gas consisting of $N$ atoms of mass $m$ is in a container that is cylindrically symmetric around the $z$-axis. Gravity $g$ acts in the negative direction of $z$. The side wall of the container is expressed as

$$z = \alpha r^\eta \tag{2.2.27}$$

in terms of the radial coordinate $r$ (as an elementary example, you may set, for example, $\eta = 2$, but the $\eta$-dependence is of some interest), where $\alpha$ is a positive constant,

---

[20] Strictly speaking, $E$ is not the internal energy, but its appropriate Legendre transform, since $\boldsymbol{B}$ is applied (a kind of enthalpy). However, here we use the usual abuse of concepts; notice that for this $E$ $dE = TdS - MdB$ instead of $dE = TdS + BdM$.

[21] In this example, $\sigma$ and the total Hamiltonian (or $h$) commute, so this simple calculation is allowed.

and the system as a whole is assumed to be thermally isolated.
(1) Find the specific heat of this system.
(2) If $\alpha$ is doubled reversible-quasistatically and adiabatically, what is the final temperature? Assume that the initial temperature is $T_0$. Also, how much work is needed to realize this process?

Let us compute the Helmholtz free energy $A$ first.[22] The canonical partition function is

$$Z = \left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2}\left[\int_0^\infty 2\pi r dr \int_{\alpha r^\eta}^\infty dz\, e^{-\beta mgz}\right]^N . \tag{2.2.28}$$

Its first term is the (momentum portion of) partition function of the ideal gas coming from the kinetic energy. The content of [ ] is[23]

$$I = \int_0^\infty \frac{1}{\beta mg} e^{-\beta mg\alpha r^\eta}\pi dr^2 = \Gamma\left(\frac{2}{\eta}+1\right)\frac{\pi}{(mg)^{1+2/\eta}\alpha^{2/\eta}}(k_BT)^{1+2/\eta}. \tag{2.2.29}$$

Internal energy may be obtained from the Gibbs-Helmholtz relation as $E = (5/2 + 2/\eta)Nk_BT$. Consequently, we have $C = (5/2 + 2/\eta)Nk_B$.

If the gas is in a cylinder (an infinitely tall cylinder), the wall is perpendicular, so $\eta = \infty$. If the container opens more widely upward, $\eta$ becomes smaller. The heat capacity becomes larger accordingly. The heat capacity increases for such a container, because increasingly more work is required to expand.

When $\alpha \to 2\alpha$ adiabatically and reversible-quasistatically, entropy stays the same. The easiest way to compute entropy is to use $(E - A)/T$:

$$S = (5/2+2/\eta)Nk_B + Nk_B \log\left[\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2}\Gamma\left(\frac{2}{\eta}+1\right)\frac{\pi}{(mg)^{1+2/\eta}\alpha^{2/\eta}}(k_BT)^{1+2/\eta}\right] \tag{2.2.30}$$

Consequently, $T^{5/2+2/\eta}/\alpha^{2/\eta}$ is kept constant. Hence, $T_f = 2^{4/(4+5\eta)}T_0$. This implies that

$$E_f = \left(\frac{5}{2}+\frac{2}{\eta}\right)Nk_B 2^{4/(4+5\eta)}T_0. \tag{2.2.31}$$

The system is thermally isolated, so the increase of internal energy must be solely due

---

[22] Precisely speaking, the potential energy due to the external field is included in the Hamiltonian in this example as well, so $A$ is actually not 'free energy,' but 'generalized Gibbs free energy' $\hat{G}$. However, in this case the external field is constant and $g$ is not regarded as a parameter that we can change during thermodynamic processes, so this distinction is not important. Therefore, notations such as $A$ and $Z$ are used informally.

[23] Since we need the exponent $Q$ such that $Z \propto \beta^Q$ to compute thermodynamic quantities, actually, we need not explicitly compute the integrals. For example, it is easy to see that the integral inside the square brackets of (2.2.28) is proportional to $\beta^{-1-2/\eta}$ by scaling the integration variables. This could even be obtained by the following dimensional-analytic argument: $[I] = [r]^2[z]$, $[z] = [r]^\eta$ and $[\beta z] = 1$, where the last relation is due to the dimensionlessness of the exponent. Thus, $z \propto \beta^{-1}$ and $r \propto \beta^{-1/\eta}$, so $I \propto \beta^{-1-2/\eta}$. In this problem, we also need the $\alpha$ dependence but that can also be obtained by analogous approaches. For example, to realize $[\alpha r^\eta] = 1$ is enough.

to work. Thus, the work needed to increase $\alpha$ is

$$W = \left(\frac{5}{2} + \frac{2}{\eta}\right) N k_B \left(2^{4/(4+5\eta)} - 1\right) T_0. \tag{2.2.32}$$

□

All the spin problems solved with the aid of the microcanonical distribution can be reworked with the aid of the canonical distribution. As we have seen for the ideal gas, Laplace-transformation with respect to $E$ often simplifies the calculation.

For the sake of completeness, some qual type questions are attached:

When the reader solves 'elementary questions,' it is a good habit to:
(o) Check that the obtained result is intuitively plausible.
(i) Check the general conclusions (e.g., consistency with thermodynamic laws).
(ii) Take various limits (e.g., $T \to \infty$).
(iii) Find approximate calculations close to a certain limit (e.g., high/low temperature analytic formulas).

**Quiz 1**. Study the system consisting of $N$ spins that can assume only up or down state. How can one study the negative temperature case? We can of course study the case with a degenerate excited state. □

**Quiz 2**. Find the magnetization of the magnetic spin as a function of temperature. □

**Quiz 3**. There is a collection of $N$ permanent electrical dipoles. When an external electric field $\boldsymbol{E}$ is applied, it must exhibit a polarization $\boldsymbol{P}$. We can compute the dielectric constant of this system. □

**Quiz 4**. The spin system problem may have the following realistic form:
Due to the ligand field the degeneracy of the $d$-orbit of the chromium ion $\mathrm{Cr}^{3+}$ is partially lifted, and the spin Hamiltonian has the following form

$$H = D(S_z^2 - S(S+1)/2), \tag{2.2.33}$$

where $D > 0$ is a constant with $S = 3/2$ (the cation is in the term ${}^4\mathrm{F}_{3/2}$).
(0) Why can we apply statistical mechanics to this 'single' ion?
(1) Obtain the occupation probability of each energy level.
(2) Calculate the entropy and the specific heat. Then, show the specific heat behaves as $\propto T^{-2}$ at higher temperatures.
(3) Suppose $C = k_B(T_0/T)^2$ with $T_0 = 0.12$K. Determine the energy spacing.
[The specific heat of this form is called the *Schottky type specific heat.*] □

**Quiz 5**. $N$ lattice sites have quantum spins of $S = 1$. Compute the spin entropy of the system. Obtain the high and low temperature asymptotic forms. □

**Quiz 6**. There is a quantum system having two states $a$ and $b$. Its Hamiltonian may be written as $H = H_0 - mB$, where $B$ is the external magnetic field, and $m$ is the magnetic moment (operator) with the following matrix elements:

$$\langle a|H_0|a\rangle = \langle b|H_0|b\rangle = E, \langle a|H_0|b\rangle = -\Delta \ (\in \boldsymbol{R}), \tag{2.2.34}$$

$$\langle a|m|a\rangle = -\langle b|m|b\rangle = \mu, \langle a|m|b\rangle = 0. \tag{2.2.35}$$

The reader may assume that $\langle a|b\rangle = 0$.
(1) When $B = 0$, obtain the canonical partition function $Tre^{-\beta H_0}$.
(2) When $B \neq 0$, obtain $\langle m\rangle$ and $\langle |m|\rangle$. Why does $\langle m\rangle$ increase with $B$? □

## 2.3 Some consequences of canonical distribution

In this section, three general consequences of the canonical distribution are discussed that may be shown easily: quantum nature of thermodynamic magnetic effects, effects of system rotation to thermodynamics, and the equipartition of energy. Canonical distribution is not indispensable, but simplifies the argument.

(1) The *Bohr-van Leeuwen theorem*r asserts that thermodynamically relevant magnetic phenomena are all quantum mechanical. The existence of a magnetic field may be expressed in terms of the vector potential $\boldsymbol{A}$ by replacing the momentum $\boldsymbol{p}$ with $\boldsymbol{p} - q\boldsymbol{A}$, where $q$ is the charge of the particle, as

$$H = \sum_i \frac{1}{2m_i}(\boldsymbol{p}_i - q_i\boldsymbol{A})^2 + V. \tag{2.3.1}$$

Here, $\boldsymbol{A}$ may be dependent on the position of the charges and includes mutual magnetic interactions due to induced charge motions.r The momentum integral can be understood as the integration over the mechanical momentum $\boldsymbol{p}-q\boldsymbol{A}$, so the classical canonical partition function cannot depend on $\boldsymbol{A}$. That is, the system free energy is independent of the magnetic field.

**Review of 4-potential**
(1) *Vector potential.* Since $div\,\boldsymbol{B} = 0$ (nonexistence of monopoles), where $\boldsymbol{B}$ is the magnetic field, we can introduce a vector field $\boldsymbol{A}$ such that $curl\,\boldsymbol{A} = \boldsymbol{B}$. This is called the *vector potential.* If $\boldsymbol{B}$ is a constant, $\boldsymbol{A} = \boldsymbol{r} \times \boldsymbol{B}/2$ may be chosen (not unique).
(2) If an electromagnetic field is given in terms of a 4-potential $(\boldsymbol{A}, \phi/c)$, $(\boldsymbol{p}-q\boldsymbol{A}, (1/c)(E-q\phi))$ is again a 4-vector, where $q$ is the charge. Its length is the same as the case without the 4-potential:

$$(1/c^2)(E - q\phi)^2 - (\boldsymbol{p} - q\boldsymbol{A})^2 = m^2c^2. \tag{2.3.2}$$

Therefore, if the speed of the particle is small $E = mc^2 + E_{classic}$ may be introduced:

$$E_{classic} = \frac{1}{2m}(\boldsymbol{p} - q\boldsymbol{A})^2 + q\phi. \tag{2.3.3}$$

(2) Equilibrium in rotating systems and non-classical rotational inertia (NCRI):**r**
We know that a system in equilibrium can have only uniform translation and rotation as its macroscopic motion. If the uniform rotational angular velocity of the system is $\boldsymbol{\omega}$, then the energy $E_{co}$ of the system seen from the 'co-rotating' coordinate system K' is (see the next fine-lettered explanation, if needed)

$$E_{co} = E - \boldsymbol{\omega} \cdot \boldsymbol{L}, \tag{2.3.4}$$

where $\boldsymbol{L}$ is the total angular momentum of the system seen from the inertial frame (lab frame) K.[24] The free energy seen from the co-rotating frame is

$$A_{co} = -k_B T \log \left[ Tr\, e^{-\beta(H - \boldsymbol{\omega} \cdot \boldsymbol{L})} \right]. \tag{2.3.5}$$

Notice that this jumppequiv2corresponds to a Legendre transformation, so

$$dA_{co} = -SdT - \langle \boldsymbol{L} \rangle \cdot d\boldsymbol{\omega} + \cdots. \tag{2.3.6}$$

The rotational moment of inertia $I$ is obtained by differentiating the angular momentum with $\boldsymbol{\omega}$:

$$I = \frac{\partial \langle \boldsymbol{L} \rangle}{\partial \boldsymbol{\omega}} = -\frac{\partial^2 A_{co}}{\partial \boldsymbol{\omega}^2}. \tag{2.3.7}$$

We may interpret this as the susceptibility of the system against changing $\boldsymbol{\omega}$.

**Example [Rotating macrobody]**
(1) Show that the rotation around the principal axis with the largest inertial moment is thermodynamically stable.
(2) Suppose a body is initially rotating with a definite total angular momentum $\boldsymbol{L}$ around an arbitrary axis through its center of mass. What happens to its temperature eventually? Assume that the system is thermally isolated.

If we could minimize the macroscopic kinetic energy $K$ of the body, since $S = S(U - K)$ , where $U$ is the total energy, we can maximize entropy. $K = L^2/2I$ implies that maximizing $I$ maximizes entropy. Therefore, the rotation around the principal axis of rotation that maximizes the moment of inertia is thermodynamically stable.[25] Since the total energy of the system must be constant, the lost rotational kinetic energy must turn into thermal energy (the temperature increases). This is actually possible through dissipation of mechanical energy due to deformation and vibration of the body.

---

[24] Henceforth, the observables with respect to the co-rotating frame will be with $'$. Actually $\boldsymbol{L} = \boldsymbol{L}'$, because $\boldsymbol{p} = \boldsymbol{p}'$ as seen below.

[25] It is also mechanically stable. The rotation is also mechanically stable around the principal axis with the minimum moment of inertia, but it is not thermodynamically stable.

**Review of mechanics seen from the rotational coordinates**

(1) Let $\boldsymbol{v}$ be the particle velocity with respect to the inertial frame K. The velocity $\boldsymbol{v}'$ seen from the co-rotating frame K' is

$$\boldsymbol{v} = \boldsymbol{v}' + \boldsymbol{\omega} \times \boldsymbol{r}. \tag{2.3.8}$$

Here, we consider the instance when K and K' coincide $\boldsymbol{r} = \boldsymbol{r}'$, and their origins are at the center of mass of the system.

(2) To identify momentum, etc., the best way is to start with the Lagrangian formalism. The Lagrangian in terms of the quantities with respect to the co-rotating frame is

$$L = \frac{1}{2}m\boldsymbol{v}'^2 + m\boldsymbol{v}' \cdot \boldsymbol{\omega} \times \boldsymbol{r}' + \frac{1}{2}m(\boldsymbol{\omega} \times \boldsymbol{r}')^2 - V. \tag{2.3.9}$$

The momentum $\boldsymbol{p}'$ is

$$\boldsymbol{p}' = \frac{\partial L}{\partial \boldsymbol{v}'} = m\boldsymbol{v}' + m\boldsymbol{\omega} \times \boldsymbol{r} = \boldsymbol{p}. \tag{2.3.10}$$

That is, the momentum seen from the inertial frame and that from the co-rotating frame are *identical*. Consequently, the angular momentum around the origin seen from the inertial frame and that from the co-rotating frame are identical as well: $\boldsymbol{L} = \boldsymbol{L}'$.

The equation of motion may be obtained with the aid of Lagrange's equation of motion:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \boldsymbol{v}'}\right) = \frac{d}{dt}(m\boldsymbol{v}' + m\boldsymbol{\omega} \times \boldsymbol{r}') = m\frac{d}{dt}\boldsymbol{v}' + m\boldsymbol{\omega} \times \boldsymbol{v}', \tag{2.3.11}$$

$$\frac{\partial L}{\partial \boldsymbol{r}'} = m\boldsymbol{v}' \times \boldsymbol{\omega} + m(\boldsymbol{\omega} \times \boldsymbol{r}') \times \boldsymbol{\omega} - \frac{\partial V}{\partial \boldsymbol{r}'}, \tag{2.3.12}$$

so we have

$$m\frac{d}{dt}\boldsymbol{v}' = -2m\boldsymbol{\omega} \times \boldsymbol{v}' - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}') - \frac{\partial V}{\partial \boldsymbol{r}'}. \tag{2.3.13}$$

The first term is the *Coriolis force*,[26] and the second the *centrifugal force*.

(3) The Hamiltonian in the co-rotating frame is defined by $H_{co} = \boldsymbol{p}' \cdot \boldsymbol{v}' - L$, so

$$H_{co} = m(\boldsymbol{v}' + \boldsymbol{\omega} \times \boldsymbol{r}) \cdot \boldsymbol{v}' - L = H - \boldsymbol{\omega} \cdot \boldsymbol{L}. \tag{2.3.14}$$

The additivity of energy and that of angular momentum imply (2.3.4) for a many-body system. Since

$$H_{co} = H - \boldsymbol{\omega} \cdot \boldsymbol{L} = \frac{1}{2}m\boldsymbol{v}'^2 + V - \frac{1}{2}m(\boldsymbol{\omega} \times \boldsymbol{r})^2, \tag{2.3.15}$$

in terms of the canonical variables

$$H_{co} = \frac{1}{2m}(\boldsymbol{p}' - m\boldsymbol{\omega} \times \boldsymbol{r})^2 + V - \frac{1}{2}m(\boldsymbol{\omega} \times \boldsymbol{r})^2. \tag{2.3.16}$$

If there are many particles, the last term becomes $(1/2)\boldsymbol{\omega}^T I_{cl}\boldsymbol{\omega}$, where $I_{cl}$ is the (classical mechanical) inertial moment tensor around the origin.

---

[26] G. G. de Coriolis (1792-1843); the Coriolis force was in 1835 [de Tocqueville, *Democracy in America* was published this year].

The Hamiltonian of a many body system seen from a rotational frame may be written as (see (2.3.16))

$$H_{co} = \sum \frac{1}{2m}(\boldsymbol{p}' - m\boldsymbol{\omega} \times \boldsymbol{r}')^2 + V - \frac{1}{2} I_{cl}\omega^2, \tag{2.3.17}$$

Therefore, if we explicitly compute the partition function in (2.3.5)

$$Z_{co} = Tr e^{-\beta H_{co}}. \tag{2.3.18}$$

Classically, just as we have done in the Bohr-van Leeuwen theorem, the $\boldsymbol{\omega}$ dependence comes only from the last term (the centrifugal potential) of (2.3.17). This implies $I = I_{cl}$. However, quantum-mechanically, this equality is not always true (e.g., if the fluid system is superfluid, then this equality breaks down). Thus, *nonclassical rotational inertia* is possible.[27]

(3) Equipartition of energy for classical systems:(**r**to ideal gas) (**r**to internal degree)**q** Let $x_i$ and $x_j$ be two components of canonical coordinates (say, the $x$-component of the spatial coordinate of particle 1 and $z$-component of the momentum of particle 2). Then, for classical systems we have[28]

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = k_B T \delta_{ij}, \tag{2.3.19}$$

where the average is over the canonical distribution. Indeed,

$$\begin{aligned} \left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle &= \frac{1}{Z}\int d\Gamma\, x_i \left[ -k_B T \frac{\partial}{\partial x_j} e^{-\beta H} \right], & (2.3.20)\\ &= -\frac{1}{Z} k_B T x_i e^{-\beta H}\Big|_{|x|\to\infty} + \frac{1}{Z} k_B T \int d\Gamma\, \frac{\partial x_i}{\partial x_j} e^{-\beta H}. & (2.3.21) \end{aligned}$$

From this we obtain the *law of equipartition of energy* for classical kinetic energy such as (no summation convention implied)

$$\left\langle \frac{p_i^2}{2m} \right\rangle = \frac{1}{2} k_B T, \tag{2.3.22}$$

---

[27] A. J. Leggett, "Can a solid be "superfluid"?", Phys. Rev. Lett., **20**, 1543 (1970) demonstrates this possibility even in solid states if there is a Bose-Einstein condensation (see the following chapter).

[28] The microcanonical version is also possible, but the demonstration is slightly cumbersome. See a problem at the end of Chapter 1.

or

$$\left\langle \frac{L_i^2}{2I_i} \right\rangle = \frac{1}{2}k_BT, \tag{2.3.23}$$

where $m$ is the mass, $I_i$ is the $i$-th principal moment of inertia ($i$-th eigenvalue of the inertial moment tensor) and $L_i$ is the corresponding component of the angular momentum. If the system is governed by a harmonic potential with a spring constant $k$, we obtain

$$\left\langle \frac{kx^2}{2} \right\rangle = \frac{1}{2}k_BT. \tag{2.3.24}$$

Since the classical kinetic energy $K$ is quadratic in (angular) momenta, Euler's theorem for homogeneous functions implies

$$\sum_i p_i \frac{\partial K}{\partial p_i} = 2K. \tag{2.3.25}$$

Thus, if there are $N$ particles, then there are $3N$ variables, so

$$\langle K \rangle = \frac{3N}{2}k_BT. \tag{2.3.26}$$

If the potential energy is a homogeneous function of $M$ position coordinates of degree $s$, then

$$s\langle V \rangle = Mk_BT. \tag{2.3.27}$$

Therefore, if a system is described as coupled harmonic oscillators with $M$ modes, the average total potential energy is[r]

$$\langle V \rangle = \frac{M}{2}k_BT. \tag{2.3.28}$$

A direct application of the equipartition of energy is the high temperature (constant volume) specific heat per particle of multiatomic molecular ideal gas. Let us assume that each molecule contains $N$ atoms. The Hamiltonian of each molecule can be written as

$$H = K_{CM} + K_{rot} + K_{vib} + U_{vib}, \tag{2.3.29}$$

where $K_X$ is the kinetic energy associated with the motion $X$: CM denotes the center of mass translational motion; rot implies rotational motion around its center of mass; vib means the vibrational motion. $U_{vib}$ is the potential energy for the vibrational motion. We may assume that the molecular internal vibrations are harmonic, so all these terms are quadratic terms. Therefore, the internal energy can be obtained only by counting the number of degrees of freedom.

The total number of degrees of freedom per molecule with $N$ atoms is $3N$.**r** For a not-linear molecule there are 3 translational degrees, and 3 rotational degrees, so there are $3N - 6$ harmonic modes. Thus, $\langle K_{CM}\rangle = \langle K_{rot}\rangle = 3k_BT/2$, and $\langle K_{vib} + U_{vib}\rangle = (3N-6)k_BT$. That is, the internal energy is $E = (3N-3)k_BT$, so $C_V = (3N-3)R$ per mole, where $R$ is the gas constant. For a linear molecule there are 3 translational degrees, and 2 rotational degrees, so there are $3N - 5$ harmonic modes.[29] Thus, $\langle K_{CM}\rangle = 3k_BT/2$, $\langle K_{rot}\rangle = k_BT$, and $\langle K_{vib} + U_{vib}\rangle = (3N-5)k_BT$. That is, $E = (3N - 5/2)k_BT$, so $C_V = (3N - 5/2)R$ per mole. It is a well-known story that these values grossly contradict the actual specific heats of molecular gases and were regarded as a paradox before the advent of quantum mechanics.

## 2.4 Entropy and information

**r**Using the canonical distribution, let us compute entropy explicitly:

$$\begin{aligned} TS &= U - A, && (2.4.1)\\ &= k_BT\log Z - k_BT\,Tr\,\frac{e^{-\beta H}}{Z}\log e^{-\beta H}, && (2.4.2)\\ &= -k_BT\,Tr\,\frac{e^{-\beta H}}{Z}\log\frac{e^{-\beta H}}{Z}. && (2.4.3)\end{aligned}$$

That is,

$$S = -k_B Tr\,\rho\log\rho, \tag{2.4.4}$$

where $\rho = e^{-\beta H}/Z$ is the canonical density operator, or, similarly, classically

$$S = -k_B\int d\Gamma\, p\log p, \tag{2.4.5}$$

where $p$ is the canonical distribution function. This is the formula first given by Gibbs in his famous book on the foundation of statistical mechanics.

The same formula was proposed by Shannon to quantify information, so (2.4.5) is often called *Shannon's formula*. It is a convenient occasion to see why such a formula

---

[29] When a molecule is straight, the reader must be able to explain into what modes the rotational degree is converted, comparing, e.g., water and carbon dioxide.

is a good measure of information. Shannon did *not* ask what information was, but tried to quantify it.[30]

Let $\eta(m)$ be the 'information' per letter we can send with a message (letter sequence) that is composed of $m$ distinct letters. Here, the word 'information' should be understood intuitively. Let us assume that all the letters are used evenly. Then, $\eta(m)$ must be an increasing function of $m$; if there are only two letters, we can send, per letter, the information telling whether $\{1,2,3\}$ or $\{4,5,6\}$ as to the outcome of a single casting of a dice, but if there are three, then more detailed information: $\{1,2\}$, $\{3,4\}$ or $\{5,6\}$ may be sent per single letter.

Now, let us use simultaneously the second set consisting of $n$ letters. We could make compound symbols by juxtaposing them as $ab$ (just as in many Chinese characters). The information carried by each compound symbol should be $\eta(mn)$, because there are $mn$ symbols. We could send the same message by sending all the left half symbols first and then the right half symbols later. The amount of information sent by these methods must be equal, so we must conclude that[31]

$$\eta(mn) = \eta(m) + \eta(n). \tag{2.4.6}$$

Since $\eta$ is an increasing function, we conclude

$$\eta(n) = c \log n, \tag{2.4.7}$$

where $c > 0$ is a constant. Its choice is equivalent to the choice of unit of information per letter and corresponds to the choice of the base of the logarithm in the formula.

If $c = 1$, we measure information in *nat*; if we choose $c = 1/\log 2$ (i.e., $\eta(n) = \log_2 n$), in *bit*. 1 bit is an amount of information one can obtain from an answer to a single yes-no question.

We have so far assumed that all the symbols are used evenly, but such uniformity is not usual. What is the most sensible generalization of (2.4.7)? We can write $\eta(n) = -\log_2(1/n)$ bits; $1/n$ is the probability for a particular letter. $-\log_2(1/n)$ may be interpreted as the expectation value of $-\log_2$(probability of a letter). This suggests that for the case with not-equal-probability occurrence of $n$ letters with probabilities $\{p_1, \cdots, p_n\}$, the expectation value of the information carried by the $i$-th symbol should be defined as $-\log_2 p_i$ bits, which is sometimes called the *surprisal*

---

[30] ⟪**Textbook of information theory**⟫ The best textbook of information theory (in English) is T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).

[31] We must send a message explaining how to combine the transferred signals as a part of the message, but the length of the needed message is finite and independent of the length of the actual message we wish to send, so in the long message limit we may ignore this overhead.

of symbol $i$, because it measures how much we are surprised by encountering this symbol (smaller $p$ should give more surprise). The average information in bits carried by a single symbol should be defined by

$$H(\{p_i\}) = -\sum_{i=1}^{n} p_i \log_2 p_i. \tag{2.4.8}$$

This is called the *Shannon information formula.*[32]**q** When Shannon arrived at (2.4.8), he asked von Neumann what it should be called. it is told that von Neumann suggested the name ‘entropy,’ adding that it was a good name because no one understood it.

**Characterization of information**
The reader may have thought the ‘derivation’ of (2.4.8) is a bit ad hoc. The formula (2.4.8) can be characterized naturally as follows. $H_n(\{p_i\})$ (in bits for $n$ letters) is:
(1) Symmetric wrt exchanging variables,
(2) $H_2(t, 1-t) = f(t)$ is positive and continuous for $t \in (0, 1)$ (this condition may be relaxed),
(3) Additive:

$$H_n(p_1, p_2, \cdots, p_n) = H_{n-1}(p_1 + p_2, \cdots, p_n) + (p_1 + p_2)H_2(p_1/(p_1 + p_2), p_2/(p_1 + p_2)). \tag{2.4.9}$$

Then, (2.4.8) is the unique form (up to the base of the logarithm).

It may be easier to use the axioms for *surprisal*. The ‘extent of surprise’ $f(p)$ we get, spotting a symbol that occurs with probability $p$, should be
(1) a monotone decreasing function of $p$ (smaller $p$ should give us bigger surprise).
(2) Nonnegative.
(3) Additive: $f(pq) = f(p) + f(q)$.[33]
Therefore, $f(p) = -c \log p$ $(c > 0)$ is the only choice. The additivity should be natural, if we consider our surprise when something rare occurs successively.

**Information per letter of English**
If the symbols of English alphabet (+ blank) appear equally probably, what is the information carried by a single symbol?(4.75 bits) In actual English sentences, it is about 1.3 bits. What actually happens? Also think about how to determine this actual value.[34]

---

[32] for an uncorrelated (or Bernoulli) information source. About Shannon himself, see S. W. Golomb et al., “Claude Elwood Shannon (1916-2002),” Notices AMS **49**, 8 (2002). His thesis is on genetics. See J. F. Crow, “Shannon’s brief foray into genetics,” Genetics **159**, 915 (2001).

[33] We could invoke the Weber-Fechner law.

[34] There is a related ‘fun’ paper, D. Benedetto, E. Caglioti and V. Loreto, “Language tree and zipping,” Phys. Rev. Lett. **88**, 048702-1 (2002). This uses the Lempel-Ziv algorithm LZ77 (used in gzip, for example) to measure information (or rather mutual information).

**Example** [Information carried by messages]
(1) Suppose a positive integer is given. It must begin with one of 1, 2, $\cdots$, and 9. If all the non-zero digits are likely to appear evenly, what is the information carried by the message that the first digit was actually 6?

Since all 9 non-zero digits are likely to appear, the initial uncertainty (entropy) is $\log_2 9 = 3.17$ bits. No uncertainty remains after receiving the message (i.e., entropy is zero), so the message must have provided the information of 3.17 bits. This is exactly the surprisal of 6 itself.

(2) In reality, it is known that the first digit does not distribute evenly. Approximately the probability that digit $D$ appears as the first digit is $P_D = \log_{10}(1 + 1/D)$. What is the information carried by this empirical law?

After knowing the law, the uncertainty (entropy) is $-\sum_D P_D \log P_D = 2.88$ bits. Therefore $3.17 - 2.88 = 0.29$ bits is the information provided by the empirical law.

(3) Now, after knowing the empirical law what is the information carried by the message that the first digit was actually 6?

With this information no uncertainty remains, so 2.88 bits must be the answer. □

In chemical physics entropy is often measured in eu (entropy unit = cal/mol·K). It may be useful to remember that 1 eu = 0.726... bits/molecule. Some people say that the unit of entropy (e.g., J/K) and unit of information (bit) are disparate. This is simply because they do not think things microscopically. If one wishes to tell each molecule to turn 'to the right', the number of required messages is comparable to the number of molecules, so it is huge, but for each molecule it is about a few bits. For example, the entropy change in a small molecule reaction is usually the order of a few eu. This is a reasonable value.

Except for the $p$ in Gibbs' formula being the probability density instead of the probability, (2.4.8) is the same as (2.4.5). Maximizing Shannon's entropy is to find the least biased distribution, so we may expect that the resultant distribution is the most probable distribution. We should be able to obtain the 'true distribution' by maximizing the Gibbs formula under the condition that we know the expectation value of energy (internal energy). This is equivalent to maximizing the following variational functional with Lagrange's multipliers $\beta$ and $\lambda$ (the latter for normalization condition):

$$-k_B \int p \log p \, d\Gamma - \beta \int pH \, d\Gamma - \lambda \int p \, d\Gamma. \tag{2.4.10}$$

This indeed gives

$$p \propto e^{-\beta H}. \tag{2.4.11}$$

The Shannon formula is derived logically from almost inescapable requirements about 'knowing something.' The above line of argument seems to indicate that the principle of statistical mechanics can be derived directly from this fundamental conceptual basis. Thus, some brave people concluded that this was the true basis of statistical mechanics; forget about mechanics, the principle of equal probability, etc.[35] This is the so-called information-theoretical derivation of statistical mechanics.

Don't be fooled by such a logic. Even if we admit that the result that maximizes the information entropy is the maximally likely result from our point of view, why does Nature have to accept it as the most 'natural' outcome? There is a logical gap here. The most natural argument to fill this gap is that we (or our brains) have evolved (or have been selected) to feel that the most natural things in the actual world are the most probable. In short, our brains have evolved in the world following the principle of equal probability. That is, the logic of information maximization is circular; implicitly, the principle of equal probability is assumed. Furthermore, if we look at (2.4.5), we should realize that something is wrong. $p$ there is not probability but probability density, so it is not invariant under coordinate transformation. For example, the description in the Cartesian coordinates and that in the equivalent polar coordinates should not give different entropies. Therefore, $\log p$ must be $\log(p/q)$ for some density distribution $q$. That is, if entropy is free from the choice of the coordinate system to describe distribution functions, the 'true' Gibbs entropy formula must read as

$$S = -k_B \int d\Gamma\, p \log \frac{p}{q}. \tag{2.4.12}$$

This is the (negative) *Kullback-Leibler entropy* whose natural implication is supplied by the large deviation theory (see just below). We cannot do anything without fixing $q$. To determine it, we need a certain statistical principle.[36]

**†Sanov's theorem**

Let $q$ be the probability density on an interval $J$. Sample $N$ points from $J$ according to this distribution, and make an empirical distribution

$$p_N = \frac{1}{N}\sum_k \delta_{x_k}, \tag{2.4.13}$$

where $\delta_y$ is the atomic measure concentrated at $y$ (i.e., $\delta(x-y)dx$, intuitively). As $N \to \infty$ $p_N$ (weakly) converges to $q$. This is the law of large numbers justifying the empirical study of distribution. It is natural to ask how quickly this convergence

---

[35] The originator seems to be E. T. Jaynes, "Information theory and statistical mechanics," Phys. Rev. **106**, 620-630 (1959).

[36] It cannot be overemphasized that even for discrete states the use of information tacitly presupposes the principle of equal probability. Think of surprisal, for example. Why is it simply a function of the probability without depending on any other contexts? It is because the world is uniform. However, this uniformity is not a consequence of any logic, but an empirical fact; we feel it natural thanks to phylogenetic learning.

occurs. This is just the large deviation question, and we can introduce the rate function(al) $I(p)$:

$$P\left(\frac{1}{N}\sum_k \delta_{x_k} \sim p\right) \approx e^{-NI(p)}. \tag{2.4.14}$$

If $x_k$ are iid stochastic variables on $J$ based on the distribution whose density is $q$, then *Sanov's theorem* asserts that

$$I(p) = \int_J dx\, p(x) \log \frac{p(x)}{q(x)}. \tag{2.4.15}$$

That is, $I(p)$ is the *Kullback-Leibler information.*[37] This implies that under the principle of equal probability Shannon's formula measures how unlikely $p$ is empirically.

It is worth remembering that (2.4.15) is non-negative for any $p$ and $q$ (as long as $q > 0$ whenever $p > 0$): We know $\log x \le x - 1$ (draw the graph to check this), so

$$-I(p) = \int_J dx\, p(x) \log \frac{q(x)}{p(x)} \le \int_J dx\, p(x) \left(\frac{q(x)}{p(x)} - 1\right) = \int_J dx\, [q(x) - p(x)] = 1 - 1 = 0. \tag{2.4.16}$$

## 2.5 Equilibrium statistical mechanics and the second law

It is often said that there are two aspects of the second law of thermodynamics:
(A) [The law about work] The work done by the system is maximum when the process is performed in a quasistatic reversible way (essentially Thomson's principle). Or, the adiabatic process without any change of the work coordinates cannot decrease the system energy (Planck's principle).
(B) [The law about irreversibility] From any initial condition any isolated system irreversibly reaches an equilibrium state.

Boltzmann (and almost all 'fundamentalists' including Einstein) tried to demonstrate (B) from mechanics and failed.[38] However, more precisely speaking, (B) is the

---

[37] There is no very kind introductory book to large deviation theory for physicists. For an informal introduction, see the introductory part and Appendix 1 of Y. Oono, "Large deviation and statistical physics," Prog. Theor. Phys., Suppl. **99**, 165-205 (1989). Sanov's theorem is demonstrated in a theoretical physicist's fashion.

[38] Boltzmann's first long paper (his second) "Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie" (Wiener Berichte **53**, 195-220 (1866)) was written when he was 22. This was one year before UIUC was chartered.

assertion of the zeroth law. As has already been noted, we must critically reconsider whether an isolated mechanical system is a meaningful idealization of a real system. It is debatable whether it is worth pursuing a demonstration of the consistency of (B) and mechanics of isolated systems. If we admit that any system eventually settles down to an equilibrium state, the content of the second law required to develop thermodynamics is only (A), so we have only to show, e.g., Thomson's principle:

$$\Delta A \leq W, \tag{2.5.1}$$

that is, the work $W$ needed to change a system from an equilibrium state $a$ to another equilibrium state $b$ with an isothermal process (i.e., a process that is realized in a system in thermal contact with a single heat bath) cannot be smaller than the free energy difference $\Delta A = A_b - A_a$, where $A_x$ is the free energy of state $x$.

Jarzynski showed the following equality (*Jarzynski's equality*) classically:[39]

$$\left\langle e^{-W/k_BT} \right\rangle = e^{-\Delta A/k_BT}. \tag{2.5.2}$$

The average on the LHS is explained in the next paragraph. The system Hamiltonian $H(\lambda)$ depends on a parameter $\lambda$ (as in Einstein's argument in Sect. 2.1) that can be changed from 0 to 1. $W$ is the work needed to change the initial system Hamiltonian $H(0)$ to the final one $H(1)$, and $\Delta A = A(1) - A(0)$, where

$$A(\lambda) = -k_BT \log Z(\lambda) \tag{2.5.3}$$

with

$$Z(\lambda) = \int d\Gamma\, e^{-\beta H(\lambda)}. \tag{2.5.4}$$

The meaning of the average on the left-hand side of (2.5.2) is the average over repetition of (i) + (ii) below:
(i) The initial condition must be sampled from the canonical distribution with the Hamiltonian $H(0)$ and temperature $T$.
(ii) While changing $\lambda$ from the initial to the final value the required work $W$ (the work we do) is measured. This measurement is done under isolation condition[40] of the system.[41]

---

[39] Jarzynski himself calls this 'irreversible work theorem.' He notes that for a cycle (i.e., $\Delta A = 0$), this was demonstrated by Bochkov and Kuzovlev; G. N. Bochkov and Yu. E. Kuzovlev, Zh. Eksp. Teor. Fiz. **72**, 238 (1977) [Sov. Phys.-JETP **45**, 125 (1977)]; Physica **106A**, 443, 480 (1981).

[40] This may be relaxed to thermally isolated condition.

[41] For the Jarzynski equality to hold, the phase space of the system must be the identical one throughout the course of experiments. Thus, free expansion of a gas to a larger volume after removing a wall (in this case no work is done, but the free energy decreases) is not an example the Jarzynski equality can cover (a remark by C. Yeung).

(iii) To make the whole process isothermal, after finishing adding work, the system is thermally connected to a heat bath of the temperature as the initial heat bath. There is no work during this last process. Eventually the system reaches an equilibrium state with the canonical distribution determined by the Hamiltonian $H(1)$.

Actually, the step (iii) has nothing to do with the measurement of $W$, so in practice only (i)+(ii) need be repeated again and again.

To demonstrate Jarzynski's equality, the expression of the work in (ii) may seem to be required, but since under the condition stated in (ii) the energy exchange is possible only in the form of work

$$W = H(y(t_F), 1) - H(y(0), 0), \tag{2.5.5}$$

where $H(y, \lambda)$ is $H(\lambda)$ with the canonical variables expressed explicitly, and $y(t)$ collectively denotes the canonical variables at time $t$.[42] Here, $t_F$ is the last time of the parameter change such that $\lambda(t_F) = 1$ (the last moment we supply work to the system). Therefore,

$$\left\langle e^{-\beta W}\right\rangle_0 = \frac{1}{Z(0)} \int d\Gamma(0)\, e^{-\beta W - \beta H(y(0),0)} = \frac{1}{Z(0)} \int d\Gamma(0)\, e^{-\beta H(y(t_F),1)}. \tag{2.5.6}$$

$d\Gamma(t)$ is the phase volume element at time $t$. Liouville's theorem holds even if the Hamiltonian is time-dependent (the phase flow is incompressible), so we may change the integration variables from $y(0)$ to $y(t_F)$ without worrying about the Jacobian: $d\Gamma(0) = d\Gamma(t_F)$. Consequently, (2.5.6) reads

$$\langle \exp(-\beta W)\rangle_0 = \frac{Z(1)}{Z(0)} = \exp[-\beta(A(1) - A(0))]. \tag{2.5.7}$$

Here, only the average over the initial ensemble shows up, but in practice it is the average over experiments starting with the sampled initial conditions.

Thanks to the convexity of the exponential function and *Jensen's inequality*:**q** for a convex function $f$, and for any well-defined averaging scheme $\langle\ \rangle$[43]

$$\langle f(x)\rangle \geq f(\langle x\rangle), \tag{2.5.8}$$

---

[42] For a $N$-particle system $y = (\boldsymbol{q}_1, \cdots, \boldsymbol{q}_N, \boldsymbol{p}_1, \cdots, \boldsymbol{p}_N)$. For example, $\dot{y}\partial/\partial y = \sum_i (\dot{\boldsymbol{q}}_i \partial/\partial \boldsymbol{q}_i + \dot{\boldsymbol{p}}_i \partial/\partial \boldsymbol{p}_i)$.

[43] **⟪Jensen's inequality⟫q** A necessary and sufficient condition for a function $f$ to be convex on a convex domain $D$ is that for any $\alpha \in [0, 1]$ and any $x_1, x_2 \in D$

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2).$$

we immediately get

$$\langle W \rangle \geq \Delta A \tag{2.5.9}$$

from (2.5.2). Sometimes, this result is interpreted as Thomson's principle; especially if the Hamiltonian returns to the original one with $\lambda = 1$, it is interpreted as Planck's principle. Indeed, the above calculation is for certain isothermal processes, so (2.5.9) is an inequality for certain isothermal processes, and has the form of Thomson's principle. However, the genuine Thomson's principle does not require any extra condition on the isothermal process. That is, the derivation of (2.5.9) only shows that equilibrium statistical mechanics does not contradict thermodynamics (in particular, the second law); it is never a derivation of the second law. If $H(0) = H(1)$, (2.5.9) gives $W \geq 0$. However, the inequality is demonstrated only for isolated systems. Planck's principle is for adiabatic processes. As already noted, thermodynamic adiabatic conditions are distinct from isolation conditions; they are quite different mechanically. Therefore, again, the inequality obtained above is, although it does not contradict thermodynamics, not the second law itself. As we will see below, all the relations connecting equilibrium statistical mechanics and the second law (A) are about isolated systems. The reason for this is trivial; mechanics cannot be used unless the system is isolated except for the parameters that may be modified systematically.

**Relevance to experiments**

The Jarzynski equality holds however fast the process may be performed. For such processes $W$ has a distribution whose peak value is far larger than $\Delta A$. Still the average of $e^{-\beta W}$ must be equal to $e^{-\beta \Delta A}$, so $W$ much smaller than $\Delta A$ must occur occasionally. That is, occasionally, the system does work upon us to compensate our work, organizing its thermal energy (or, we can say, occasionally producing a large amount of negative entropy). Therefore, although the inequality is correct even in the thermodynamic limit, we must be extremely patient to accumulate very rare fluctuation results. In other words, verifying this equality for a macrosystem is in principle impossible. The correctness of the second law implies the impossibility of the demonstration of Jarzynski's equality for macroscopic systems. Consequently, the equality, if meaningful at all, may be practical only for mesoscale systems (e.g., single macromolecules) at the largest.

Then, it is very inconvenient to study isolated systems, because molecules such as proteins are usually immersed in a solution. To cover this situation, we take the system (e.g., a molecule) + the solution (environment) as an isolated system just as we have derived the canonical ensemble from the microcanonical ensemble. The total

---

From this it is trivial to obtain for $\alpha_i \in [0,1]$ ($\sum \alpha_i = 1$) and for $x_i \in D$

$$f\left(\sum \alpha_i x_i\right) \leq \sum \alpha_i f(x_i).$$

Thus, the resultant inequality for countably many points and the continuity of the convex function give Jensen's inequality.

Hamiltonian reads

$$H(y(t), \lambda(t)) = H_s(z(t), \lambda(t)) + H_{int}(y(t)) + H_0(x(t)), \tag{2.5.10}$$

where $z$ collectively denote the canonical variables of the system, $x$ collectively denote the canonical variables of the solution (environment), $y = (z, x)$, $H_s$ is the system Hamiltonian, $H_0$ the Hamiltonian of the environment and $H_{int}$ is the interaction Hamiltonian between the system and the environment. We have assumed that the controllable parameter $\lambda$ is attached only to the system Hamiltonian. We sample the initial condition of the whole system according to the canonical distribution of the whole system at temperature $T$ and then change $\lambda$ just as the isolated case discussed already. We obtain exactly as (2.5.7)

$$\langle \exp(-\beta W(t)) \rangle = \frac{Y(t)}{Y(0)}, \tag{2.5.11}$$

where

$$Y(t) = \int d\Gamma(x) d\Gamma(z) \exp[-\beta(H_s(z, \lambda(t)) + H_{int}(y) + H_0(x))]. \tag{2.5.12}$$

In contrast to the derivation of the canonical distribution, we cannot ignore the interaction term, because the system is small. Notice that the effective canonical partition function may be written as

$$Z^*(t) = Z_0 \int d\Gamma(z) \, \langle \exp[-\beta(H_s(z, \lambda(t)) + H_{int}(z, x))] \rangle_0 = Z_0 e^{-\beta A^*(t)}, \tag{2.5.13}$$

where $\langle \ \rangle_0$ is the average over the canonical distribution over the environment, $Z_0$ is the canonical partition function for the environment, and $A^*$ is the thermodynamic potential of the system immersed in the environment. Obviously, $Y(t)/Y(0) = Z^*(t)/Z^*(0)$, so with this careful definition of the thermodynamic potential, (2.5.11) reads

$$\langle \exp(-\beta W) \rangle = \exp(-\beta \Delta A^*). \tag{2.5.14}$$

Notice that there is no formal difference from (2.5.2).

Next, let us consider the quantum case of Jarzynski's equality. Consider a quantum system with the Hamiltonian $H$. In terms of kets, the Schrödinger equation for the system reads

$$i\hbar \frac{d}{dt} | \ \rangle = H | \ \rangle. \tag{2.5.15}$$

Taking the conjugate of this equation, we obtain the Schrödinger equation for bras:

$$i\hbar \frac{d}{dt} \langle \ | = -\langle \ | H. \tag{2.5.16}$$

If the initial ket is given by $|0\rangle$, the solution $|t\rangle$ of (2.5.15) at time $t$ is obtained as

$$|t\rangle = e^{-iHt/\hbar}|0\rangle. \tag{2.5.17}$$

$e^{-iHt/\hbar}$ is the time evolution operator for the system, and is a unitary operator. Generally speaking, a canonical transformation in classical mechanics corresponds to a unitary transformation in quantum mechanics. Even if the system is perturbed by an external disturbance or manipulation, this is true. The Hamiltonian in (2.5.15) can now be an arbitrary function of $t$, but the relation between the initial ket and the ket at time $t$ is connected by a unitary transformation $U(t)$ as

$$|t\rangle = U(t)|0\rangle. \tag{2.5.18}$$

Let us demonstrate that $U(t)$ is indeed unitary. Since

$$i\hbar\frac{dU}{dt} = H(t)U, \tag{2.5.19}$$

we immediately see $U^*U = 1$. To show $UU^* = 1$ is easy, if $H$ does not explicitly depend on $t$; more generally, we use $U(t+\Delta t) = (1+H(t)\Delta t)U(t)+o[\Delta t]$ to see

$$U(t+\Delta t)U^*(t+\Delta t) = (1+H(t)\Delta t/i\hbar)U(t)U^*(t)(1-H(t)\Delta t/i\hbar)+o[\Delta t]. \tag{2.5.20}$$

Then, we can use mathematical induction: if $U(t)U(t)^* = 1$ holds up to time $t$,

$$U(t+\Delta t)U^*(t+\Delta t) = (1+H(t)\Delta t/i\hbar)(1-H(t)\Delta t/i\hbar) = 1+o[\Delta t]. \tag{2.5.21}$$

Since $U(0) = 1$, for all $t > 0$ we have shown that $U(t)$ is unitary.

A density operator $\rho$ is a bilinear form of bras and kets (a linear combination of the products of a ket and a bra), so from (2.5.18) and (2.5.19) the time evolution of the density operator of the system whose Hamiltonian is $H$ is governed by the following *von Neumann equation*:[44]

$$i\hbar\frac{d}{dt}\rho = H(t)\rho - \rho H(t) = [H(t), \rho]. \tag{2.5.22}$$

---

[44] The time dependence of the density operator can be determined by $Tr\ \rho A(t) = Tr\rho(t)A$, where $A$ is an observable, and $A(t)$ is $A$ in the Heisenberg's picture. That is, $\rho(t)$ corresponds to $\rho$ in the Schrödinger picture.

$$i\hbar\frac{d}{dt}Tr\{\rho A(t)\} = Tr\{\rho[A(t), H]\} = Tr\{[H, \rho(t)]A\} = Tr\left\{i\hbar\frac{d\rho}{dt}A\right\}$$

for any $A$, so we must conclude (2.5.22).

Its solution to the initial value problem may be written as

$$\rho(t) = U(t)\rho(0)U^*(t). \tag{2.5.23}$$

Let us consider a quantum system with a Hamiltonian $H(\lambda)$ dependent on a parameter $\lambda$. The initial parameter value is $\lambda = 0$ and the final one is $\lambda = 1$. The eigenvalues of $H(\lambda)$ are denoted as $E_n(\lambda)$ and the corresponding eigenket is $|n;\lambda\rangle$. The initial density operator is the canonical density operator $\rho(0) = e^{\beta A(0) - \beta H(0)}$, where $A(0)$ is the initial free energy. Let $U$ be the time evolution operator corresponding to the parameter change from $\lambda = 0$ to 1 under the condition that the system is isolated. As we have already noted, if the mechanical work $W$ is identified with the difference between the final and the initial Hamiltonians (e.g., $W = U^*H(1)U - H(0)$), then the noncommutativity of $H(1)$ and $H(0)$ generally ruins the equality. However, if we identify $W$ as the difference between the observed final and the initial energies $E_n(1) - E_m(0)$, then the Jarzynski equality in the following form holds:

$$\left\langle e^{-\beta(E_n(1)-E_m(0))} \right\rangle = e^{-\beta \Delta A}, \tag{2.5.24}$$

where the average is over the initial states $m$ and final states $n$ (reached from state $m$) and $\Delta A = A(1) - A(0)$. To write this explicitly, we need the transition probability between the two observation results $E_m(0)$ initially and $E_n(1)$ finally. This can be written as $P_{n \leftarrow m} = |\langle n;1|U|m;0\rangle|^2$. Notice that $P_{n\leftarrow m}$ is doubly stochastic: $\sum_n P_{n\leftarrow m} = \sum_m P_{n\leftarrow m} = 1$.[45] Explicitly, the left hand side reads

$$\sum_{n,m} e^{-\beta(E_n(1)-E_m(0))} P_{n\leftarrow m} e^{\beta(A(0)-E_m(0))} = \sum_{nm} e^{-\beta E_n(1)} P_{n\leftarrow m} e^{\beta A(0)} = e^{-\beta\Delta A}. \tag{2.5.25}$$

Therefore, Jensen's inequality tells us the inequality $W \geq \Delta A$ compatible with the second law. However, as noted already in the classical case, it is not the second law itself.

Lenard essentially demonstrated the following statement quantum mechanically:[46]

| | | |
|---|---|---|
| The second law of thermodynamics | $\Longleftrightarrow$ | The equilibrium distribution is a monotone decreasing function of energy. |

If $\rho(0) = \sum_n |n;0\rangle w(E_n(0))\langle n;0|$ with $w(x)$ being a decreasing function of $x$, then

$$\langle H\rangle_U - \langle H\rangle_0 \geq \langle H\rangle_{ad} - \langle H\rangle_0, \tag{2.5.26}$$

---

[45] $\{P_{i\leftarrow j}\}$ is a stochastic matrix, if $P_{i\leftarrow j} \geq 0$ and $\sum_i P_{i\leftarrow j} = 1$. It is called *doubly stochastic*, if $\sum_j P_{i\leftarrow j} = 1$ also holds.

[46] A. Lenard, "Thermodynamical proof of the Gibbs formula for elementary quantum systems," J. Statist. Phys. **19**, 575 (1978).

where $\langle\ \rangle_0$ is the expectation value with respect to the initial distribution $\rho(0)$, and $\langle\ \rangle_U$ is the average over $U\rho(0)U^*$ for an arbitrary unitary operator $U$, and $\langle\ \rangle_{ad}$ implies the average over the distribution reached from $\rho(0)$ by a quantum mechanical adiabatic process. If the initial and the final Hamiltonians are identical (2.5.26) becomes Planck's principle *under* isolation condition.

**†An outline of the proof of (2.5.26)**
Let us assume that eigenvalues are ordered as $E_i(\lambda) \le E_{i+1}(\lambda)$. Let the initial distribution be $\rho_0 = \sum_n |n;0\rangle w(E_n(0))\langle n;0|$. We have only to show that if $w(x)$ is nonincreasing as a function of $x$, for any unitary transformation $U$

$$Tr(HU\rho_0U^{-1}) \ge \langle H\rangle_{ad}. \tag{2.5.27}$$

In terms of components this reads

$$\sum_{nm} E_n(1)P_{nm}w(E_m(0)) \ge \sum_n E_n(1)w(E_n(0)). \tag{2.5.28}$$

where $P_{nm}$ is the doubly stochastic matrix already given in (2.5.25). Therefore, if we can show the following inequality, we are done: Let $A_i$ and $B_i$ be both increasing sequences. Then, for any doubly stochastic matrix $P_{ij}$

$$\sum_{ij} P_{ij}A_iB_j \le \sum_i A_iB_i. \tag{2.5.29}$$

To prove this we have only to find an operation $Q$ such that $A^TPB \le A^T(QP)B$ and $Q^nP = I$ for some finite positive $n$.[47]. Let $P$ be an $N \times N$ matrix with nonzero elements $P_{aN}$ and $P_{Nb}$ for some $a, b$ ($a, b \in \{1, \cdots, N\}$; if there is no such pair replace $N$ with $N-1$ and repeat the argument here). $Q$ is a procedure to find such a pair $a, b$ and change the following four elements $P_{aN}, P_{Nb}, P_{NN}, P_{ab}$ as follows. Let the smaller one of $P_{aN}$ and $P_{Nb}$ be $\Delta$: $\Delta = \min\{P_{aN}, P_{Nb}\}$:

$$\begin{aligned}(QP)_{aN} &= P_{aN} - \Delta, &(2.5.30)\\ (QP)_{Nb} &= P_{Nb} - \Delta, &(2.5.31)\\ (QP)_{ab} &= P_{ab} + \Delta, &(2.5.32)\\ (QP)_{NN} &= P_{NN} + \Delta. &(2.5.33)\end{aligned}$$

Then,

$$A^T(QP)B - A^TPB = (A_a - A_N)(B_b - B_N)\Delta \ge 0. \tag{2.5.34}$$

$QP$ remains to be a doubly stochastic matrix, and at least one element in the $N$-th raw or column goes from a positive value to zero. Repeat applying $Q$ until there is no pair $a, b$ such that $P_{aN}$ and $P_{Nb}$ are simultaneously nonzero. Within finitely many application of $Q$ all the elements of the $N$-th raw and column become zero except for $P_{NN}$ $(= 1)$. Thus, the problem has been reduced to a $N-1 \times N-1$ problem. Since the theorem is trivial for $N = 1$, we are done.

---

[47] The proof here is due to Nir Friedman (private commun., 2008).

Conversely, the second law implies that $w(x)$ must be decreasing. If the fourth law of thermodynamics is also assumed, $w(E) \propto e^{-\beta E}$ follows with some 'technical' conditions. More precisely, the assumptions are:
(i) *Passivity*: This is essentially the second law. Let $H(t)$ be a time-dependent Hamiltonian and the density operator $\rho(t)$ obeys von Neumann's equation

$$i\hbar \frac{d\rho}{dt} = [H(t), \rho]. \tag{2.5.35}$$

Here, we assume that this Hamiltonian has no time dependence outside the range of time $[0, 1]$ and $H(0) = H(1)$. Then, the work needed for this cycle is nonnegative:

$$W(K, \rho_0) = \int_0^1 dt\, Tr \rho(t) \frac{dH(t)}{dt} \geq 0. \tag{2.5.36}$$

(ii) Structural stability: any small perturbation of the Hamiltonian does not destroy the system passivity.
(iii) Let a compound system consist of two subsystems and be in a passive structurally stable state. Its density operator is the product[48] of the density operators of the two subsystems.

Roughly speaking, (i) implies that the distribution is a monotone decreasing function of energy, and (iii) restricts the functional form to be exponential. Consequently, the canonical distribution is derived. If we do not demand (iii), more general distributions could be obtained.

Lenard's equivalence relation sounds plausible, even if the isolation condition is replaced with the thermodynamic adiabatic condition. However, thermodynamic adiabatic conditions lack mechanical interpretation, such a thermodynamically meaningful assertion can never be proved by mechanics. Furthermore, the mechanical interpretation of work follows the tradition initiated by Einstein, and pays no attention to whether it is macroscopically realizable or not. In short, we do not have any understanding of heat in terms of mechanics, so no satisfactory relation between the second law and statistical mechanics is obtained.

---

[48] More precisely, tensor product

## 2.6 Various ensembles

In this section, we introduce various convenient partition functions. We have already encountered two distributions (ensembles), microcanonical and canonical. Let us observe their relations and corresponding thermodynamic potentials to see some general patterns. Here, notations are the same as before.[49]

$$w(E) \quad \rightarrow \quad Z(T) = \int dE\, w(E) e^{-\beta E} = \int dE\, e^{[S(E)-E/T]/k_B}, \tag{2.6.1}$$

$$S(E) = k_B \log w(E) \quad \rightarrow \quad S(E) - \frac{E}{T} = -\frac{A}{T} = k_B \log Z(T). \tag{2.6.2}$$

The reader must have realized a certain formal relation. The Laplace transformation of the partition function in (2.6.1) corresponds to the Legendre transformation of the thermodynamic potential. That is, the Laplace transformation of the partition function and the Legendre transformation of the thermodynamic potential correspond one to one. Let us first see that this is a general feature in classical statistical mechanics.

As we see above, it is convenient to use the Gibbs relation with respect to entropy:

$$dS = \frac{1}{T} dE + \frac{P}{T} dV - \frac{\mu}{T} dN - \frac{x}{T} dX. \tag{2.6.3}$$

If we wish to study a system under constant temperature and pressure, the convenient *Massieu function*[50] must be $S - E/T - PV/T = -A/T - PV/T = -G/T$. Therefore, following the above pair of formulas, we have[q]

$$Z(T, V) \quad \rightarrow \quad Q(T, P) = \int dV\, Z(T, V) e^{-\beta PV}, \tag{2.6.4}$$

$$-\frac{A}{T} = k_B \log Z(T, V) \quad \rightarrow \quad -\frac{A}{T} - \frac{PV}{T} = -\frac{G}{T} = k_B \log Q(T, P). \tag{2.6.5}$$

The ensemble describing a system under constant temperature and pressure is called a *pressure ensemble.*

If we wish to study a system under constant temperature and volume but attached

---

[49] As noted in Section 1.5 we can use $W(E)$ (the number of all the microstates whose energy does not exceed $E$) instead of $w(E)$ (the number of microstates in the energy shell $(E - \delta E, E]$). Here, we use $w(E)$ because it is the logical choice.

[50] The Legendre transforms of entropy are generally called *Massieu functions.*[r]

to a chemostat allowing the change of the number of particles $N$ (a summary of chemical thermodynamics is in Section 1.12), then the convenient Massieu function is

$$q = S - \frac{E}{T} + \frac{\mu N}{T} = \frac{PV}{T}, \tag{2.6.6}$$

which is sometimes called *Kramers' q*. We have

$$Z(T,N) \quad \to \quad \Xi(T,\mu) = \int dV\, Z(T,N) e^{\beta\mu N} = \sum_N e^{[-A/T+\mu N/T]/k_B} \tag{2.6.7}$$

$$-\frac{A}{T} = k_B \log Z(T,N) \quad \to \quad -\frac{A}{T} + \frac{\mu N}{T} = q = k_B \log \Xi(T,\mu). \tag{2.6.8}$$

$\Xi$ is called the *grand canonical partition function*.

Let us summarize this Laplace-Legendre correspondence:r

$$\begin{tikzcd}
\Theta \arrow[r, leftrightarrow, "\Theta = \log \Upsilon(x)"] \arrow[d, "\text{Legendre}"] & \Upsilon(x) \arrow[d, "\text{Laplace}"] \\
\Lambda = \sup_x(\Theta + xX) \arrow[r, leftrightarrow, "\Lambda = \log \Gamma(X)"] & \Gamma(X) = \int dx\, \Upsilon(x) e^{\beta x X} = \int dX\, e^{\Theta(X) + \beta x X}
\end{tikzcd}$$

The above general theory is correct within the classical statistical mechanics framework. In quantum statistical mechanics, if $H$ and $X$ are not commutative, generally we respect the inequality: $Tr\, e^{-\beta(H-xX)} \neq Tr\, e^{-\beta H} e^{\beta x X}$, so the formulas cannot generally be as simple as the above, because exponential factors cannot be successively multiplied. However, for the usual systems the Hamiltonian and the particle number are commutative, so the above commutative diagram is true for the relations among microcanonical, canonical and grand canonical partition functions even for quantum statistical mechanics.

**Example [Ideal rubber band]** (cf microcanonical approach)
There is a polymer consisting of $N$ monomers of length $a$. Each monomer can orient in any direction freely (if there is no external tensile force).
(1) We wish to use temperature $T$ and tensile force $F$ as independent thermodynamic variables to describe the system. What generalized canonical partition function $Y$ should we use? Relate $Y$ to $Z(T,L)$ (or define $Y$ in terms of $Z$) that is the canonical partition function for a polymer with one end fixed at the origin and the other end fixed at $L$ along the axis parallel to the tensile force to be applied).
(2) Compute $Y$ as a function of $F, T$ and $N$. Choose the force direction to be the polar axis direction so that the azimuthal angle may be used to describe the tilt of the monomer from the force direction.

(3) Find the end-to-end distance $L$ of the polymer as a function of $F$.
(4) For a small force the polymer is a harmonic chain. Find the Hooke's constant.
(5) Obtain the mean square fluctuation of the end to end distance (i.e., $\langle \delta L^2 \rangle$) without the external force under constant $T$.

The Legendre transformation we wish to use is $A \to \tilde{G} = A - FL$ (or $-A/T \to -A/T + FL/T$), so

$$Y(T,F) = \sum_L Z(T,L)e^{\beta FL} = \int dL\, e^{\beta FL} Z(T,L). \tag{2.6.9}$$

Notice that[51]

$$LF = Fa \sum_i \cos\theta_i. \tag{2.6.10}$$

Thus,

$$Y = \int \prod_i (d\theta_i d\varphi_i \sin\theta_i) \exp\left(\beta Fa \sum \cos\theta_i\right) = \prod_i \left[\int d\theta_i d\varphi_i \sin\theta_i \exp\left(\beta Fa \cos\theta_i\right)\right], \tag{2.6.11}$$

$$= \left[2\pi \int_0^\pi d\theta \sin\theta \exp\left(\beta Fa\cos\theta\right)\right]^N = \left[2\pi \int_{-1}^1 dx\, e^{\beta Fax}\right]^N. \tag{2.6.12}$$

That is,

$$Y(T,F) = \left[4\pi \frac{\sinh(\beta aF)}{\beta Fa}\right]^N. \tag{2.6.13}$$

We should realize that $Z$ is much harder to compute.

Thermodynamics tells us that $\tilde{G} = -k_BT\log Y$ ($d\tilde{G} = -SdT - LdF$), so

$$L = +k_BT \left.\frac{\partial \log Y}{\partial F}\right|_T = Na\Lambda(\beta aF), \tag{2.6.14}$$

where $\Lambda(x)$ is called the *Langevin function* and is defined by

$$\Lambda(x) = \coth x - \frac{1}{x}. \tag{2.6.15}$$

The length of the polymer should not exceed $Na$; in the $F \to \infty$ limit $\Lambda \to 1$, so indeed $L \to Na$.

If the external force is small, we may expand as

$$\Lambda(x) = \frac{1}{3}x - \frac{1}{45}x^3 + \cdots. \tag{2.6.16}$$

Therefore, Hooke's constant is $k = 3k_BT/Na^2$. This model is a model of rubber elasticity. $k \propto T$ indicates the entropic nature of the elasticity. Notice that $a^2N/3 =$

---

[51] As noted in Section 1.5 $Z(T,L)$ should have not only the portion explicitly used here but also that due to thermal motion of monomers, but here the latter is irrelevant, so it is suppressed.

$\langle R^2\rangle$ is the variation of the end-to-end distance (called the mean square end-to-end distance). We will learn its close relation to the spring constant shortly (see the fluctuation-response relation).

□

There are different partition functions, but how are their results related? From the above formal consideration, it appears all give us the same thermodynamics; thermodynamic potentials corresponding to various partition functions are related mutually by Legendre transformations. Therefore, for systems with ordinary particle-particle interactions (e.g., not long-ranged and with a sufficiently hard core),[52] in the thermodynamic limit all the partition functions are expected to be equivalent in the sense that all give identical thermodynamics.**r** Let us demonstrate this *ensemble equivalence***r** with a representative example: the equivalence of microcanonical and canonical ensembles. The relation between the canonical and the grand canonical partition functions is an exercise at the end of this Chapter.[53]

Let us take a finite system with $N$ particles. The Laplace transformation may be understood as a summation over energy shells (this was the original form in Section 2.1; the ground state energy is set to be zero):

$$Z(T)=\sum_E w(E)e^{-\beta E}, \tag{2.6.17}$$

where the sum $\sum_E$ is the sum over all the shells of thickness $\delta E$. $w(E)$ is the number of microstates whose energy is in $(E-\delta E, E]$. (2.6.17) is an infinite sum, but we assume it converges (if not, there is no thermodynamics[54]). Therefore, there is $E_0$ (which is of order $N$ as shown in the fine lettered explanation below) beyond which the sum (2.6.17) is bounded by 1:

$$\sum_E w(E)e^{-\beta E}=Z(T)\le \sum_{E\le E_0} w(E)e^{-\beta E}+1. \tag{2.6.18}$$

Let us write the largest value of the summand of (2.6.18) as $w(E^*)e^{-\beta E^*}$ ($\ge w(0)\ge 1$).**r** Obviously (see Fig. 2.6.1),

$$w(E^*)e^{-\beta E^*}\le Z(T)\le (E_0/\delta E)w(E^*)e^{-\beta E^*}+1\le 2(E_0/\delta E)w(E^*)e^{-\beta E^*} \tag{2.6.19}$$

---

[52] So long as entropy is a concave function ($-S$ is a convex function) of all the extensive parameters = thermodynamic coordinates.

[53] As has already been noted, the relations among these three ensembles are exactly the same for classical and quantum systems.

[54] Its convergence is guaranteed if the interactions among particles are as mentioned above.

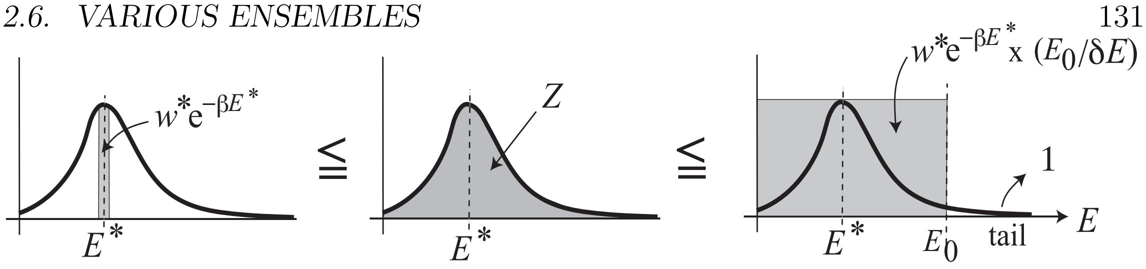


Fig. 2.6.1 The inequalities in (2.6.19) are illustrated. The gray areas respectively from left to right correspond to the three formulas in (2.6.19). $E^*$ is the peak position, whose height is $w^* e^{-\beta E^*}$, where $w^* = w(E^*)$.

Therefore,

$$\sup_E (S(E) - E/T) \leq -A/T \leq \sup_E (S(E) - E/T) + k_B \log(2E_0) - k_B \log \delta E. \tag{2.6.20}$$

Since $E_0$ is of order $N$ and since we know $\langle E \rangle / N \ll \delta E \ll \langle E \rangle$, the above formula implies that the free energy per particle $A/N$ obtained from the canonical partition function, that is obtained with the aid of Laplace transformation from the microcanonical partition function, and the free energy per particle obtained thermodynamically with the aid of Legendre transformation from the entropy, that is the result of the microcanonical partition function, are identical. This is the equivalence of ensembles. The demonstration given here is almost rigorous. The only insufficient part is that it is not demonstrated that $E_0$ can be chosen to be of order $N$ (see the following detailed explanation).

**Outline of demonstration of $E_0 = O[N]$**
We need an upper bound of $w(E)e^{-\beta E}$. Obviously, $w(E) \leq W(E)$, so we estimate an upper bound of $W(E)e^{-\beta E}$.

The demo consists of two steps. First, $W$ is estimated for non-interacting systems, and then we consider how the result is modified by interactions. For non-interacting systems (with suffix $I$) it is not hard to show[55] that the general form is:

$$W^I(E) = \exp[V s(e)], \tag{2.6.21}$$

---

[55] ⟪**Upper bounds for ideal systems, an outline**⟫ For the classical case, we already know an explicit estimate. For quantum cases no explicit microcanonical calculation is given, but for fermion systems, it is clear that any microstate is also a microstate of the corresponding classical system, so $W$ is bounded by the classical result. For boson systems with the aid of Kronecker's deltas we can write (here $\varepsilon_i$ is the energy of the $i$-th one particle state, and $n_i$ is the number of particles having this energy, the occupation number of the $i$-th state; see Chapter 3)

$$W(E) = \sum_{E' \leq E} \sum_{\{n_i\}} \delta_{\sum n_i \varepsilon_i, E'} \delta_{\sum n_i, N}.$$

Then, noting the following obvious inequality for any $\beta$ and $\mu$

$$\delta_{\sum n_i \varepsilon_i, E'} \delta_{\sum n_i, N} \leq e^{\beta[E' - \sum n_i \varepsilon_i] + \beta\mu[N - \sum n_i]} \leq e^{\beta[E - \sum n_i \varepsilon_i] + \beta\mu[N - \sum n_i]}$$

where $e = E/V$ is the energy density. $s$ is the entropy density, so it is convex upward. Since $\partial s/\partial e = 1/T$, as long as $\beta_0 > 0$, there is $e = e_0$ such that $\partial s/\partial e = \beta_0$. Thus, we have

$$s(e) \le s(e_0) + \beta_0(e - e_0). \tag{2.6.22}$$

Consequently, we have

$$W^I(E) \le \exp\{V[s(e_0) + \beta_0(e - e_0)]\}. \tag{2.6.23}$$

Let us arrange all the energy eigenvalues of the system (resp., of the corresponding non-interacting system) in the increasing order as $E_i \le E_{i+1}$ (resp., $E_i^I \le E_{i+1}^I$). If the interaction is stable, that is, if the potential energy $\Phi \ge -BN$ ($B > 0$), there is an operator inequality $H \ge H^I - BN$, where $H$ is the system Hamiltonian, and $H^I$ is that for the corresponding non-interacting system (i.e., $H = H^I + \Phi$). Then, with the aid of the *minimax principle* (see the next entry), we can generally show

$$E_i \ge E_i^I - BN. \tag{2.6.24}$$

This implies the following inequality:

$$W(E) \le W^I(E + BN) \le \exp\{V[s(e_0 + B\rho) + \beta_0(e + B\rho - e_0)]\} = \exp[V\tilde{s} + \beta_0 E], \tag{2.6.25}$$

where $e = E/V$, $\rho = N/V$ and $\tilde{s} = s(e_0 + B\rho) + \beta_0(B\rho - e_0)$.

Hence,

$$W(E)e^{-\beta E} \le \exp[V\tilde{s} + (\beta_0 - \beta)E], \tag{2.6.26}$$

but $\beta_0$ can be anything if positive, so, for example, we may choose $\beta_0 = \beta/2$:

$$W(E)e^{-\beta E} \le \exp[V\tilde{s} - \beta E/2]. \tag{2.6.27}$$

Thus,

$$\sum_{E \ge E_0} w(E)e^{-\beta E} \le \sum_{E \ge E_0} W(E)e^{-\beta E} \le \sum_{E \ge E_0} \exp[V\tilde{s} - \beta E/2] = \frac{2}{\beta} e^{V\tilde{s}} e^{-\beta E_0/2}. \tag{2.6.28}$$

We wish to bound this with 1, so we require

$$\frac{\beta E_0}{2} + \log\frac{\beta}{2} > V\tilde{s}. \tag{2.6.29}$$

That is, $E_0 \ge (2k_BT)[V\tilde{s} + \log(2k_BT)]$ is required. Therefore, when $T$ is given, if the volume is sufficiently large, we have only to choose $E_0/V$ slightly larger than $2k_BT\tilde{s}$. Indeed, $E_0 = O[V] = O[N]$.

---

we have

$$W(E) \le e^{\beta[E - \mu N]}\Xi(1/\beta, \mu),$$

where $\Xi$ is the grand canonical partition function for the corresponding classical system, which can be estimated easily. See Chapter 3. Notice that $E$, $N$ and $\log\Xi$ are extensive, so we have the form (2.6.21).

**The minimax principle for eigenvalues**[56]

**Theorem** Let $A$ be a self-adjoint operator defined on a vector space $\mathcal{H}$. For an arbitrary finite dimensional subspace $\mathcal{M}$ of $\mathcal{H}$, let us compute

$$\lambda(\mathcal{M}) = \sup_{|\varphi\rangle \in \mathcal{M}, \langle\varphi|\varphi\rangle=1} \langle\varphi|A|\varphi\rangle. \tag{2.6.30}$$

Let us arrange the eigenvalues of $A$ in the increasing order and number them (with their multiplicity taken into account) and write the $k$-th eigenvalue as $\mu(k)$. Then, we have

$$\mu(k) = \lambda(k) \equiv \inf_{dim\mathcal{M}=k} \lambda(\mathcal{M}). \tag{2.6.31}$$

That is, if we choose an arbitrary $k$-dimensional subspace $\mathcal{M}$ and compute (2.6.30), then the smallest among them is the $k$th eigenvalue of $A$.

[Demo] Let us write the orthonormal basis corresponding to $\{\mu(k)\}$ as $\{|k\rangle\}$.

First of all, $\lambda(k) \leq \mu(k)$ is obvious: if we adopt the $k$-dimensional subspace spanned by $\{|1\rangle, \cdots, |k\rangle\}$ as $\mathcal{M}$, $\lambda(\mathcal{M}) = \mu(k)$, so the smallest value we look for by changing $\mathcal{M}$ cannot be larger than this value.

Thus, we have only to show $\lambda(k) \geq \mu(k)$. Since $\mathcal{M}$ is a finite dimensional vector space, it must be contained in a subspace spanned by $\{|1\rangle, \cdots, |N\rangle\}$, if $N$ is sufficiently large. Take a subspace $V$ spanned by $\{|k\rangle, \cdots, |N\rangle\}$. Since $\mathcal{M}$ is with dimension $k$, $\mathcal{M}$ and $V$ must share a vector which is not zero. Let us normalize it and call it $|0\rangle$. By the definition of $\lambda(\mathcal{M})$

$$\lambda(\mathcal{M}) \geq \langle 0|A|0\rangle, \tag{2.6.32}$$

but $|0\rangle = \sum_{i=k}^{N} |i\rangle\langle i|0\rangle$, so

$$\langle 0|A|0\rangle = \sum_{i=k}^{N} |\langle i|0\rangle|^2 \mu_i \geq \mu(k). \tag{2.6.33}$$

That is, for any $k$-dimensional $\mathcal{M}$ we have $\lambda(\mathcal{M}) \geq \mu(k)$. Thus, we have shown that $\lambda(k) = \mu(k)$. □

Let $A$ and $B$ be self-adjoint and $A \geq B$, that is, for any vector $\varphi$ $\langle\varphi|A|\varphi\rangle \geq \langle\varphi|B|\varphi\rangle$. Let us arrange the eigenvalues of $A$ (respectively, the eigenvalues of $B$) in increasing order (with the multiplicity due to degeneracy taken into account) and write the $k$th eigenvalues as $\lambda_A(k)$ (respectively, $\lambda_B(k)$). Then, for any $k$, the inequality is preserved:

$$\lambda_A(k) \geq \lambda_B(k). \tag{2.6.34}$$

This is the relation we needed above, which can be shown with the aid of the minimax principle. Notice that for any $k$ dimensional subspace $\mathcal{M}$ there is a vector $\phi$ satisfying[57]

$$\langle\phi|A|\phi\rangle \geq \lambda_B(\mathcal{M}). \tag{2.6.35}$$

---

[56] D. Ruelle, *Statistical Mechanics* (World Scientific, 1999; original 1969) Section 2.5. Here, the demonstration is made slightly informal.

[57] Here, 'sup' is used instead of 'max,' so some care is needed. A precise statement is: for any

Therefore, $\lambda_A(\mathcal{M}) \geq \lambda_B(\mathcal{M})$. Next, we choose $\mathcal{M}$ to minimize $\lambda_A(\mathcal{M})$. Then, with a similar argument, we must conclude that there is a $k$ dimensional subspace $\mathcal{M}'$ satisfying

$$\lambda_A(k) \geq \lambda_B(\mathcal{M}'). \tag{2.6.36}$$

Hence, we are done. □

The ensemble equivalence allows us to use any convenient partition function we wish in order to compute thermodynamic quantities.[58] As long as the interaction is not crazy,[59] the equivalence is valid.[60]

If the system under study is small (say, $N = 300$), then the ensemble equivalence is not right. In such a case, a practical recommendation is to use the ensemble whose condition is close to the actual system (e.g., if under constant temperature and pressure, use the pressure ensemble corresponding to $G$). The system Hamiltonian may have to include the system-environment interactions explicitly (see below).

An intuitive picture of a statistical ensemble is a collection of non-interacting macroscopic systems. That this picture may be used when we discuss thermodynamic observables of a single macroscopic system is a key observation to set up statistical mechanics as discussed in Section 1.1. Needless to say, a real macroscopic system is not at all a collection of independent subsystems. However, there are cases for which the system may be literally understood as an ensemble of independent subsystems. For such systems, the microstates of the whole collections obey the principle of equal probability.[61] In this case, the principle of equal probability governs the true

---

$\varepsilon > 0$ there is a vector $\phi$ satisfying

$$\langle \phi | A | \phi \rangle \geq \lambda_B(\mathcal{M}) - \varepsilon.$$

Therefore, $\lambda_A(\mathcal{M}) \geq \lambda_B(\mathcal{M}) - \varepsilon$. Hence, $\lambda_A(\mathcal{M}) \geq \lambda_B(\mathcal{M})$.

[58] This statement is unconditionally correct for classical statistical mechanics. For quantum statistical mechanics, microcanonical, canonical and grand canonical ensembles are equivalent. For other ensembles some care is needed [see Supplementary Pages].

[59] As long as entropy is concave (convex *upward*), our logic goes through.

[60] **⟪Phase coexistence and Legendre transformation⟫** If entropy is convex upward (i.e., if the system is thermodynamically normal), then the equivalence relation holds. However, as we will see in Sect. 5.13, states with distinct $\{E, X_i\}$ may be mapped to an identical state $\{T, X_i\}$ (many to one correspondence). Even in such cases, thermodynamic variables unique in both ensembles agree. For example, the temperature of the state $\{T, X_i\}$ agrees with the temperature calculated from any of $\{E, X_i\}$.

[61] **⟪Asymptotic equipartition property⟫** Let us assume that the microstate $i$ of each independent element is realized with probability $p_i$. A microstate of the whole system is designated

distribution of the microstates of the collection. Therefore, the distribution we can obtain as a canonical distribution applied to a collection of independent elements is the actual distribution for such a collection attached to a heat bath. Utilizing the fact that each element is statistically independent (and the Hamiltonian of the whole system $H$ is the sum of Hamiltonians $h_i$ of individual elements as $H = \sum h_i$), we can get the statistical description of the individual system as the marginal distribution obtained by projection onto each element as

$$e^{-\beta H} = e^{-\beta(h_1+h_2+\cdots+h_N)} \to e^{-\beta h_i}. \tag{2.6.37}$$

That is, we can apply the canonical distribution to individual elements. We know the simplest such example from elementary statistical mechanics: the $\mu$-space description of an ideal gas molecule (the *Maxwell distribution*): the probability of a molecule with mass $m$ to have the velocity $\boldsymbol{v}$ is proportional to $\exp(-m|\boldsymbol{v}|^2/2k_BT)$. A far more interesting example is a dilute solution of macromolecules. In this case we may regard the system as an ensemble of droplets containing a single macromolecule.

When we study a single mesoscopic system (e.g., a molecular motor sitting on a polystyrene bead), if we can observe it repeatedly with a sufficient time interval (much longer than the system relaxation time[62]), then the collection of the observed data may be understood as those obtained by observing an ensemble of noninteracting mesoscopic systems. If we use the Hamiltonian of the mesoscopic system with *interaction terms* with its environment, the usual canonical formalism may be used.[63] Perhaps, in certain cases we wish to use an ensemble under constant tensile force. In any case, if the the Hamiltonian describing the mesoscopic system contains all the interaction terms with its environment, we can fully utilize the formalism explained in this section.

---

by specifying the microstates of individual elements $\{i_k\}_{k=1}^N$, where $N$ is the number of independent elements in the collection. The weak law of large numbers tells us $-(1/N)\log P(\{i_k\}_{k=1}^N) = -(1/N)\sum_{k=1}^N \log p_{i_k} \to -\sum p_i \log p_i = h$ ($h$, which is the Shannon entropy for each system). The strong law also holds: for almost all the microstates of the whole system $P(\{i_k\}_{k=1}^N) \simeq e^{-nh}$, which is called the asymptotic equipartition property.

[62] *Relaxation time* is a measure of memory retention time of the system.

[63] Perhaps one may say with constant pressure, but in condensed phases the system compressibility is so low that usually we need not worry about the pressure effect.

## 2.7 Fluctuation and response

So far we have studied very large systems (thermodynamic limit). There, the law of large numbers reigns: for any $\varepsilon > 0$

$$P\left(\left|\frac{1}{V}X(V) - \langle X\rangle\right| > \varepsilon\right) \to 0 \tag{2.7.1}$$

in the $V \to \infty$ limit, where $V$ is the volume of the domain we observe, and $X(V)$ is the observed result there.[64] If we observe a volume that is not large, then, the probability above is positive. That is, we observe fluctuations. Even if we say the volume we observe is tiny, since we are macroscopic organisms, the volume is sufficiently large from the microscopic point of view. Therefore, fluctuations should not be very large, and we have only to consider the second moments to quantify fluctuations.[65]

The most important observation we will make is that quantities with large fluctuations can easily be perturbed externally. There is a close relation between the system response and fluctuations in the system.

Take a finite (classical) system and observe a work coordinate $X$ there. We assume that the system is maintained at temperature $T$. Let us look at the response of $X$ to the modification of its conjugate variable $x$ (with respect to energy). A convenient partition function is the following generalized canonical partition function:

$$Z(\beta, x) = Tre^{-\beta H + \beta x\hat{X}}, \tag{2.7.2}$$

where $\hat{X}$ is the microscopic description of the work coordinate $X$ $(= \langle\hat{X}\rangle)$. We obtain[66]

$$d(\log Z(\beta, x)) = -Ed\beta + Xd(\beta x), \tag{2.7.3}$$

so the susceptibility $\chi = \partial X/\partial x$ of the response $X$ to the change of $x$ reads[67]

$$\chi = \beta\frac{\partial^2 \log Z}{\partial(\beta x)^2}. \tag{2.7.4}$$

---

[64] If $X$ is extensive, $X(V)$ is the total amount in $V$ (i.e., $X(V)/V$ is its density). If $X$ is intensive, then $X(V)/V$ should be interpreted as the average value in the volume $V$.

[65] If the reader remembers large deviation theory, she will realize that we are discussing the quadratic approximation of the rate function.

[66] In the following, the system interacts with a system preparing the external field $x$, and the interaction potential energy is $-x\hat{X}$. The canonical distribution based on this + the intrinsic Hamiltonian of the system is a generalized canonical distribution. The corresponding thermodynamic potential is $A - xX$. See the discussion on thermally isolated systems.

[67] In the following partial derivative, the variables kept fixed depends on the situation. In this case $X$ can fluctuates, but whether other variable $Y$ can fluctuate or not depends on the situation.

Performing the partial differentiation, we obtain**r**

$$\chi = \beta\left(\langle \hat{X}^2\rangle - \langle \hat{X}\rangle^2\right) = \beta\langle \delta\hat{X}^2\rangle \geq 0, \tag{2.7.5}$$

where $\delta\hat{X} = \hat{X} - \langle\hat{X}\rangle$. This is the *fluctuation-response relation.* (**r**to rotating system, **r**to spring constant, **r**to Gaussian distribution, **r**to critical exponent)

We can make three important observations from the result.
(i) The 'ease' of response results from 'large' fluctuations. Notice that $\chi$ describes the response to an external perturbation, but the variance of $\hat{X}$ is due to spontaneous thermal fluctuations. Gentle nudging of the system (reversible change) must respect the spontaneity of the system.
(ii) Since $X$ is extensive and $x$ is intensive, $\chi$ must be extensive (proportional to the number of particles there, $N$)**r**. Therefore, $\delta X = O[\sqrt{V}]$.[68]
(iii) $\chi$ cannot be negative. This is the manifestation of the stability of the equilibrium state (free energy minimum) as we will see in more detail later.

If there are many variables fluctuating simultaneously, we should extend the above result $\delta X = \beta\langle\delta\hat{X}^2\rangle\delta x$ to a collection $\{X_i\}$ of extensive variables. A natural guess may be

$$\delta X_i = \sum_j \beta\langle\delta\hat{X}_i\delta\hat{X}_j\rangle\delta x_j, \tag{2.7.6}$$

where $\delta X_i$ is the thermodynamically observable perturbation result when conjugate variables $\{x_i\}$ are perturbed. This formula can be obtained most easily by the Taylor expansion of the following formula wrt $\{\delta x_i\}$:[69]

$$\begin{aligned}
\langle X_i\rangle &= \int d\Gamma\,\hat{X}_i e^{-\beta H(x_j+\delta x_j)}\bigg/\int d\Gamma\, e^{-\beta H(x_j+\delta x_j)} && (2.7.7)\\
&= \int d\Gamma\,\hat{X}_i e^{-\beta\left[H(x_j)-\sum_j X_j\delta x_j+\cdots\right]}\bigg/\int d\Gamma\, e^{-\beta\left[H(x_j)-\sum_j X_j\delta x_j\cdots\right]} && (2.7.8)\\
&= \langle\hat{X}_i\rangle_0 + \sum_j \beta\left(\langle\hat{X}_i\hat{X}_j\rangle_0 - \langle\hat{X}_i\rangle_0\langle\hat{X}_j\rangle_0\right)\delta x_j + \cdots. && (2.7.9)
\end{aligned}$$

---

For example, if a magnet is thermostatted and can freely change its volume, then $T, P$ are the variables that must be kept constant, but if the volume is not allowed to change, $T, V$ must be fixed.

[68] away from critical points. There, $\chi$ can diverge (fluctuation diverges)**r**, so nothing can be said from this argument.

[69] Here, $H$ is the Hamiltonian including the intereaction between the system and the external field.

Here $\langle\ \rangle_0$ implies the average without perturbation $\{\delta x_j\}$. In the RHS of (2.7.6), $\langle\ \rangle$ is just the equilibrium average, so it is actually $\langle\ \rangle_0$ in the above result. (2.7.6) may be written as**rq**

$$\langle \delta\hat{X}_i \delta\hat{X}_j \rangle = k_B T \left( \frac{\partial X_i}{\partial x_j} \right)_{x_1 \cdots \check{x}_j \cdots x_n}, \tag{2.7.10}$$

where $\check{\ }$ implies to remove $x_j$. Notice that the Maxwell relations follow from this. In the above formulas, $X$ and $\hat{X}$ are meticulously distinguished, but if not confusion is likely, we may simply drop hats.

The reader should have realized that the derivation of (2.7.6) does not work for quantum mechanical cases, because Taylor expansion is not simple due to noncommutativity of the variables (observables); since generally $e^{A+B} \neq e^A e^B$, we cannot write $e^{A+\epsilon B} = e^A(1+\epsilon B+\cdots)$. To obtain the quantum version, we need the following non-commutative Taylor expansion formula:

$$e^{A+a} = e^A \left( 1 + \int_0^1 d\lambda\, e^{-\lambda A} a e^{\lambda A} + \cdots \right). \tag{2.7.11}$$

†A formal proof is left to standard textbooks.[70] Here, it is explained how to guess this formula, using *Trotter's formula*[71]

$$e^{A+a} = \lim_{n\to\infty} \left( e^{A/n} e^{a/n} \right)^n. \tag{2.7.12}$$

From this, we obtain

$$\left. \frac{d}{dx} e^{A+xa} \right|_{x=0} \simeq \sum_{j=1}^{n} \frac{1}{n} e^{A(1-j/n)} a e^{A(j/n)}. \tag{2.7.13}$$

This is a Riemann sum formally converging to the integral in the above formula.

With the aid of (2.7.11) we can Taylor expand as

$$e^{-\beta(A+a)} = e^{-\beta A} \left( 1 + k_B T \int_0^\beta d\lambda\, e^{\lambda A} a e^{-\lambda A} + \cdots \right). \tag{2.7.14}$$

---

[70] One way is to show:

$$\frac{d}{d\lambda} e^{\lambda(A+a)} e^{-\lambda A} = e^{-\lambda(A+a)} a e^{-x\lambda A}.$$

Integrating this from $t = 0$ to 1, we obtain the desired formula.

[71] ⟪**Trotter's formula**⟫ This formula holds if at least $A$ or $a$ is a bounded operator. See J. Glimm and A. Jaffe, *Quantum Mechanics, a functional integral point of view*, second edition (Springer, 1987), Section 3.2. Its ('infinite' matrix) component representation is the path integral. The idea of Trotter's formula is very important in the quantum Monte Carlo method and in numerical solutions of (nonlinear) partial differential equations. The corresponding formula for finite matrix $A$ is originally due to Lie, so the formula should be called the *Lie-Trotter formula*.

If we introduce the Heisenberg picture $A(t) = e^{iHt/\hbar}Ae^{-iHt/\hbar}$,

$$e^{\beta H}Ae^{-\beta H} = A(-i\beta\hbar), \tag{2.7.15}$$

so

$$e^{-\beta(A+a)} = e^{-\beta A}\left(1 + k_BT\int_0^\beta d\lambda\, a(-i\lambda\hbar) + \cdots\right). \tag{2.7.16}$$

Using this and repeating the calculation just as the classical case, we arrive at

$$\delta X_i = \sum_j \beta\langle\delta\hat{X}_i;\delta\hat{X}_j\rangle\delta x_j, \tag{2.7.17}$$

where

$$\langle\hat{X};\hat{Y}\rangle = k_BT\int_0^\beta d\lambda\,\langle\hat{X}(-i\lambda\hbar)\hat{Y}\rangle_e \tag{2.7.18}$$

is called the *canonical correlation function*. Here, the suffix $e$ implies the equilibrium average without perturbation. If all the variables commute, this reduces to $\langle\hat{X}\hat{Y}\rangle$ and (2.7.17) reduces to (2.7.6). It is often the case that if we replace the correlation in classical results with the canonical correlation, we obtain formulas valid in the quantum cases.

## 2.8 Thermodynamic fluctuation theory and system stability

Fluctuation is a window to the smaller scales of the world supporting what we observe. Even in equilibrium fluctuations are time-dependent, so fluctuation allows us to glimpse nonequilibrium behaviors as well. Here, we study only equal time correlations among various fluctuating quantities.[72]

Einstein in 1910[73] studied the deviation of thermodynamic observables in a small domain of a system from their equilibrium values in order to understand critical

---

[72] The best reference is Landau-Lifshitz, *Course of Theoretical Physics*, volume 5 Chapter XII (Pergamon Press 1980). Note that $k_B$ is not explicitly written in this book.

[73] [This year, Russel and Whitehead, *Principia Mathematica* started to publish (∼1913), the Mexican revolution began, Tolstoy died (11/20, 1817∼) Nightingale died.]

fluctuations.[74] To obtain the probability of fluctuations, he inverted the Boltzmann principle as

$$w(\{X\}) = e^{S(\{X\})/k_B}, \tag{2.8.1}$$

where $\{X\}$ collectively denotes extensive variables. Then, he postulated that the statistical weight for the value of $X$ deviated from its equilibrium value may also be obtained by (2.8.1). Since we know the statistical weights, we can compute the probability of observing $\{X\}$ as

$$P(\{X\}) = \frac{w(\{X\})}{\sum_{\{X\}} w(\{X\})}. \tag{2.8.2}$$

The denominator may be replaced with the largest term in the summands (cf. Sect. 2.6), so we may rewrite the formula as

$$P(\{X\}) \simeq \frac{w(\{X\})}{w(\{X_{eq}\})} = e^{[S(\{X\})-S(\{X_{eq}\})]/k_B} = e^{-|\Delta S|/k_B}, \tag{2.8.3}$$

where $\simeq$ implies the equality up to a certain unimportant numerical coefficient, and $\{X_{eq}\}$ is the value of $\{X\}$ that gives the largest $w$ (maximizes the entropy), that is, the equilibrium value. $\Delta S = S(\{X\}) - S(\{X_{eq}\})$ is written as $-|\Delta S|$ to emphasize the sign of $\Delta S$ (i.e., negative). To the second order

$$P(\{\delta X\}) \propto e^{-|\delta^2 S|/k_B}. \tag{2.8.4}$$

Einstein proposed this as the *fundamental formula for small fluctuations* in a small portion of *any* equilibrium system.[75] Here, $S(\{X_{eq}\})$ is the genuine thermodynamic entropy as a thermodynamic function, and is a univalent function in the thermodynamic space. In contrast, in $S(\{X\})$ $X$ involves deviations from equilibrium ($\delta X = X - X_{eq}$ is the deviation), so $S$ is not the entropy as a thermodynamic function. In the Einstein theory, it is assume that even for this, the standard thermodynamic relation is applicable. That is, if $\delta X$ is not too large, $S(X) - S(X_{eq})$ is approximated as a quadratic form of $\{\delta X\}$ whose coefficients may be computed with the aid of equilibrium thermodynamics. This is a fundamental postulate of thermodynamic fluctuation theory.

---

[74] A. Einstein, "Theorie der Opaleszenz von homogenen Flüssigkeitsgemischen in der Nahe des kritischen Zustandes," Ann. Phys., **33**, 1275-1298 (1910). [Theory of critical opalescence of homogeneous fluid mixture near the critical state]. J. D. Jackson, *Classical Electrodynamics*, 2nd Edition (Wiley, 1975) Sect. 9.7 is a good summary of related topics.

[75] The large deviation theory**r** is the theory of fluctuation. $|\Delta S|/k_B$ divided by the small volume being observed is the rate function.

**†Einstein is ‘always correct’**

The reader surely claims that Einstein cannot be right. $w$ is a good quantity to study only isolated systems, so if we wish to study a small portion embedded in a large system, the proper formalism must be

$$P(\{\delta X\}) = \frac{\hat{Z}(\{X\})}{\hat{Z}(\{X_{eq}\})} = e^{-|\delta\Psi|/k_B} \tag{2.8.5}$$

instead of (2.8.2), where $\hat{Z}$ is an appropriate generalized canonical partition function we have already discussed:

$$\Psi = k_B \log \hat{Z} \tag{2.8.6}$$

is the corresponding Massieu function. It is an appropriate Legendre transformation of $S$ as

$$\Psi = S + \sum_j \frac{\hat{x}_j X_j}{T}, \tag{2.8.7}$$

where the intensive variables fixed externally are hatted. The reader is correct, but Einstein is also (almost and practically always) correct.

Let us choose $X_i$ as independent variables. Then,

$$\Psi = \Psi_{eq} + \delta^2\Psi + \cdots = \left[S_{eq} - \sum_j \frac{x_j \delta X_j}{T} + \delta^2 S + \cdots\right] + \sum_j \frac{\hat{x}_j \delta X_j}{T} + \cdots. \tag{2.8.8}$$

Under an equilibrium condition $x_i = \hat{x}_i$ must hold. That is, *to the second order* the answer agrees with Einstein irrespective of the starting ensemble. In practice we never compute higher order correlation than the second order, so Einstein is ‘always’ right. We can recover the classical statistical case discussed in the preceding section from the above formalism. Notice, however, that the first variation $\delta S$ that Einstein conveniently ignored is not zero. Still, the second variation term of an appropriate thermodynamic potential is identical to $\delta^2 S$.

In contrast, the thermodynamic fluctuation theory in this section imagines a mesoscopic region in a large system and consider fluctuations of various quantities there. Therefore, there is no variable that is kept constant (except, for example, for the situation we consider fluctuations in a fixed small fictitious volume), so any thermodynamic variable can fluctuate. Still, to the second order we may always use $\delta^2 S$.

Then, what variables should we use as independent variables to compute the second variation? Suppose we study a system that requires $n$ thermodynamic coordinates (i.e., its thermodynamic space is $n$-dimensional). To change statistical ensembles, extensive variables $X_i, \cdots$ are replaced with their conjugate intensive variable $x_i, \cdots$. We know any ensemble may be used to study the second moments of thermodynamic fluctuations. As we see from this, generally speaking, we may choose $n$ independent variables, selecting one (i.e., $X$ or $x$) from each conjugate pair $\{x, X\}$. Any choice will do, but sometimes a clever choice may drastically simplify the calculation.

To study the fluctuation we need the second order variation $\delta^2 S$. This can be computed from the Gibbs relation (here $\delta$ means the so-called ‘virtual variation,’ but notice that such variations are actually spontaneously realized by thermal fluctuations)

$$\delta S = \frac{1}{T}(\delta E + P\delta V - \mu\delta N - x\delta X) \tag{2.8.9}$$

as (this is the second order term of the Taylor expansion, so do not forget the overall factor 1/2)[76]

$$\begin{aligned}
\delta^2 S &= \frac{1}{2}\left[\delta\left(\frac{1}{T}\right)(\delta E + P\delta V - \mu\delta N - x\delta X) + \frac{1}{T}(\delta P\delta V - \delta\mu\delta N - \delta x\delta X)\right], \quad (2.8.10)\\
&= -\frac{\delta T}{2T^2}T\delta S + \frac{1}{2T}(\delta P\delta V - \delta\mu\delta N - \delta x\delta X). \quad (2.8.11)
\end{aligned}$$

Thus, we have arrived at the following useful expression worth remembering (actually, almost nothing to remember anew):[77]

$$\delta^2 S = \frac{-\delta T\delta S + \delta P\delta V - \delta\mu\delta N - \delta x\delta X}{2T} = -\frac{1}{T}\delta^2 E.\mathbf{r} \tag{2.8.12}$$

Consequently, the probability density of fluctuation can have the following form, the starting point of practical calculation of fluctuations (second moments):

$$P(\text{fluctuation}) \propto \exp\left\{-\frac{1}{2k_BT}(\delta T\delta S - \delta P\delta V + \delta\mu\delta N + \delta x\delta X)\right\}. \tag{2.8.13}$$

---

[76] If the reader has some trouble in understanding the following formulas, look at a simple example: $f = f(x, y)$, where $x$ and $y$ are regarded as independent variables. If we can write

$$\delta f = X\delta x + Y\delta y,$$

then

$$\delta X = \frac{\partial X}{\partial x}\delta x + \frac{\partial X}{\partial y}\delta y, \quad \delta Y = \frac{\partial Y}{\partial x}\delta x + \frac{\partial Y}{\partial y}\delta y.$$

Therefore, the second order Taylor expansion term reads

$$\delta^2 f = \frac{1}{2}\left(\frac{\partial X}{\partial x}\delta x^2 + \frac{\partial X}{\partial y}\delta y\delta x + \frac{\partial Y}{\partial x}\delta x\delta y + \frac{\partial Y}{\partial y}\delta y^2\right) = \frac{1}{2}(\delta X\delta x + \delta Y\delta y).$$

In short, the second variations of independent variables are zero (i.e., $\delta^2 x = \delta^2 y = 0$):

$$\delta[X\delta x + Y\delta y] = \delta X\delta x + X\delta^2 x + \delta Y\delta y + Y\delta^2 y = \delta X\delta x + \delta Y\delta y.$$

[77] As has already been stated in the above fine lettered explanation notice that $\delta S = 0$ does not hold. The derivation of this formula by Einstein was indeed a feat.

$\delta^2 E$ is the energy we must supply, if we wish to create the fluctuation by ourselves. Therefore, we may rewrite (2.8.13) as**q**

$$P(\text{fluctuation}) \propto e^{-\beta W_f}, \tag{2.8.14}$$

where $W_f$ is the reversible work required to create the fluctuation.

The formula in the round parentheses of (2.8.13) is a quadratic form in independent variations (of our choice, say, $\{\delta T, \delta P, \delta N, \delta X\}$).[78] Thus, this is a *multivariate Gaussian distribution*.

**Multivariate Gaussian distribution**
A multivariate distribution is called the *Gaussian* distribution, if any marginal distribution is Gaussian. Or more practically, we could say that if the negative log of the density distribution function is a positive definite quadratic form (apart from a constant term due to normalization) of the deviations from the expectation values, it is Gaussian:

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{det(2\pi V)}} \exp\left(-\frac{1}{2}\boldsymbol{x}^T V^{-1}\boldsymbol{x}\right), \tag{2.8.15}$$

where $V$ is the *covariance matrix* defined by (do not forget that our vectors are column vectors)

$$V = \langle \boldsymbol{x}\boldsymbol{x}^T \rangle. \tag{2.8.16}$$

If the mean value of $\boldsymbol{x}$ is nonzero $\langle \boldsymbol{x} \rangle = \boldsymbol{m}$, simply replace $\boldsymbol{x}$ in the above with $\boldsymbol{x} - \boldsymbol{m}$. The reader must not have any difficulty in demonstrating that (2.8.15) is correctly normalized.

In particular, for the two variable case:**q**

$$f(x,y) \propto \exp\left\{-\frac{1}{2}\left(ax^2 + 2bxy + cy^2\right)\right\}, \tag{2.8.17}$$

then

$$V = \Lambda^{-1} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}^{-1} = \frac{1}{det\,\Lambda}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \tag{2.8.18}$$

That is,

$$\langle x^2 \rangle = c/det\,\Lambda, \quad \langle xy \rangle = -b/det\,\Lambda, \quad \langle y^2 \rangle = a/det\,\Lambda. \tag{2.8.19}$$

The fluctuation-response relation (classical case) can be obtained with the aid of thermodynamic fluctuation theory as well. If $\delta^2 E$ is expressed in terms of $\delta X$ (deviations of extensive variables), (2.8.13) generally reads

$$P(\text{fluctuation}) \propto \exp\left\{-\frac{1}{2k_B T}\sum_{i,j}\left(\frac{\partial x_i}{\partial X_j}\right)_{x_1\cdots\check{x}_j\cdots x_n}\delta X_i \delta X_j\right\}. \tag{2.8.20}$$

---

[78] We can freely choose any one variable from each of the intensive-extensive pairs as already discussed in a fine-lettered explanation.

Here, ˇ implies to remove the variable below the check mark. Notice that

$$\left[Matr.\left(\left.\frac{\partial x_i}{\partial X_j}\right|_{X_1\cdots\check{X}_j\cdots X_n}\right)\right]^{-1}=Matr.\left(\left.\frac{\partial X_i}{\partial x_j}\right|_{x_1\cdots\check{x}_j\cdots x_n}\right). \tag{2.8.21}$$

Therefore, we obtain the fluctuation-response relation:

$$\langle\delta X_i\delta X_j\rangle=\beta\left.\frac{\partial X_i}{\partial x_j}\right|_{x_1\cdots\check{x}_j\cdots x_n}. \tag{2.8.22}$$

This thermodynamic formalism can be applied if $\delta^2 S$ is well defined irrespective of quantum or classical nature of the fluctuations. If the fluctuation is too rapid (i.e., $T\Delta S\sim h/\tau$ for the fluctuation time scale $\tau$[79]), or the temperature is too low, we cannot rely on the theory. That is, for (2.8.4) to be applicable, we need the following condition:

$$\Delta S\gg h/T\tau. \tag{2.8.23}$$

Our remaining task is to use practically the fundamental formula (2.8.13). Notice that
(i) If an extensive variable $X$ and an intensive variable $y$ are not conjugate (with respect to energy), then $\langle\delta X\delta y\rangle=0$. Therefore, if one are interested in only $\langle\delta X^2\rangle$, one should choose other variables than $\delta X$ all intensive, so that no cross terms show up (as seen in the example just below).
(ii) $\langle\delta X_i\delta x_i\rangle=k_BT$ (no summation convention).
(iii) The fluctuations of the quantities that may be interpreted as the expectation values of microscopically-mechanically expressible quantities (e.g., internal energy, volume, pressure), the fluctuations tend to zero in the $T\to 0$ limit. However, the fluctuations of entropy and the quantities obtained by differentiating this do not satisfy the above property.

All the above statements require demonstration (good exercises).

Calculate $\langle\delta T^2\rangle$. Let us assume there is no fluctuation in the number of particles ($\delta N=0$). To this end, we must first choose independent variables. $\delta T$ must of course be chosen. We need one more independent variable. Referring to (i), we should choose $\delta V$ as the other variable. Then, we have only to compute the following term in (2.8.13):

$$-\frac{1}{2k_BT}\delta S\delta T=-\frac{1}{2k_BT}\left.\frac{\partial S}{\partial T}\right|_V\delta T^2+\cdots=-\frac{C_V}{2k_BT^2}\delta T^2+\cdots. \tag{2.8.24}$$

---

[79] due to the energy-time uncertainty principle. This is the relation between the certainty of the energy of a state and its relaxation time scale.

Therefore, we can easily conclude

$$\langle \delta T^2 \rangle = k_B T^2 / C_V. \tag{2.8.25}$$

The same result should be obtained even if our choice of the independent variables is not clever.

Let us compute the same $\langle \delta T^2 \rangle$ with $\delta T$ and $\delta P$ as independent fluctuations. Then,

$$-\frac{1}{2T}[\delta S \delta T - \delta P \delta V] = -\frac{1}{2T}\left\{ \left.\frac{\partial S}{\partial T}\right|_P \delta T^2 - 2\left.\frac{\partial V}{\partial T}\right|_P \delta T \delta P - \left.\frac{\partial V}{\partial P}\right|_T \delta P^2 \right\}. \tag{2.8.26}$$

Using (2.8.19), we should be able to compute the desired quantity. Since

$$det\,\Lambda = \frac{1}{T^2}\frac{\partial(V,S)}{\partial(T,P)}, \tag{2.8.27}$$

$$\langle \delta T^2 \rangle = -\frac{k_B}{det\,\Lambda}\frac{1}{T}\left.\frac{\partial V}{\partial P}\right|_T = -k_B T \frac{\frac{\partial(V,T)}{\partial(P,T)}}{\frac{\partial(V,S)}{\partial(T,P)}} = k_B T \frac{\partial(V,T)}{\partial(V,S)} = \frac{k_B T^2}{C_V}. \tag{2.8.28}$$

We know $\delta^2 S < 0$ for equilibrium states, so there is no too large spontaneous fluctuations. This is an indication of the stability of the equilibrium state.**r** (**r**to Sect 2.1, **r**to Sect 2.7) We could also write $\delta^2 E > 0$. Generally speaking, for any virtual changes

$$\delta^2 S < 0 \text{ or } \delta^2 E > 0 \tag{2.8.29}$$

is called the *stability criterion* for the equilibrium state. We can also call

$$\delta^2 S > 0 \text{ or } \delta^2 E < 0 \tag{2.8.30}$$

the *evolution criterion*; if this inequality holds, the state must change (evolve). (Notice that these criteria do not depend on the ensemble choice. Why?) Here, remember that the variation $\delta^2 S$ (sometimes called virtual displacement) is *actually* being checked incessantly and spontaneously by the system with the aid of thermal fluctuations.

The stability criterion to the second order reads

$$\sum_{i,j}\frac{\partial^2 E}{\partial X_i \partial X_j}\delta X_i \delta X_j > 0. \tag{2.8.31}$$

That is, $\delta^2 E$ must be a positive definite quadratic form of $\{\delta X_i\}$. A necessary and sufficient condition for a square matrix $A$ to give a positive definite quadratic form is that all the principal minors of $A$ is positive. In particular, all the diagonal elements must be positive:[r]

$$\left.\frac{\partial x_i}{\partial X_i}\right|_{X_1\cdots\check{X}_j\cdots X_n} > 0. \tag{2.8.32}$$

This inequality is *Le Chatelier's principle*:[80] the change due to perturbation of the system occurs in the direction to reduce the perturbation effect. For example, if we put heat into a system, then the system temperature goes up, so further import of heat becomes difficult. Such a relation is thanks to the convexity of $E$, so differentiability of $E$ (or $S$) is not needed.[81] Therefore, we can conclude, for example,[q]
(i) if a phase I changes into II by raising temperature, then the entropy of phase II must be larger than that of I.
(ii) if a phase I changes into II by increasing pressure, then the volume of II must be smaller than that of I (do not forget that the conjugate variable of $P$ is not $V$ but $-V$).

> In $^3$He at low temperatures, melting of solid occurs upon decreasing temperature. Therefore, (i) implies that the solid state has a larger entropy (due to nuclear spin disorder[82]) than liquid (*Pomeranchuk effect*).

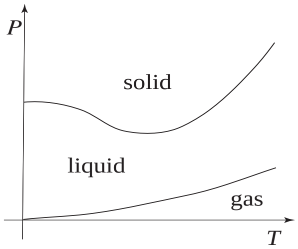


Fig. 2.8.1 A schematic phase diagram of $^3$He. Under constant pressure at low temperatures (below ~ 0.3 K), heating solidifies $^3$He liquid. Therefore, entropy increases upon solidification according to (i). In this case the solid phase is always above the liquid phase, so the volume must shrink upon solidification according to (ii) even if melting is caused by heating.

The positivity of $2 \times 2$ principal minors implies

$$\frac{\partial(x_i, x_j)}{\partial(X_i, X_j)} > 0. \tag{2.8.33}$$

We have already encountered this matrix in the calculation of the two variable fluctuation above.

The positivity of the diagonal terms may be used to derive the following inequality.

---

[80] Henry Louis Le Chatelier (1850-1936).

[81] $E$ is $C^1$ as a function of extensive variables, but it may not be twice differentiable. Thus, (2.8.32) may not be meaningful. Still, for example, $-\Delta P \Delta V > 0$ must hold.

[82] These spins order into an antiferromagnetic state only at around $10^{-3}$ K.

Since $dx = \frac{\partial x}{\partial X}\big|_Y dX + \frac{\partial x}{\partial Y}\big|_X dY + \cdots$

$$\begin{aligned}
\left.\frac{\partial x}{\partial X}\right|_y &= \left.\frac{\partial x}{\partial X}\right|_Y + \left.\frac{\partial x}{\partial Y}\right|_X \left.\frac{\partial Y}{\partial X}\right|_y, && (2.8.34)\\
&= \left.\frac{\partial x}{\partial X}\right|_Y + \left.\frac{\partial x}{\partial Y}\right|_X \frac{\partial(Y,y)}{\partial(X,x)}\frac{\partial(X,x)}{\partial(X,Y)}\frac{\partial(X,Y)}{\partial(X,y)} && (2.8.35)\\
&= \left.\frac{\partial x}{\partial X}\right|_Y - \left.\frac{\partial x}{\partial Y}\right|_X^2 \left.\frac{\partial Y}{\partial y}\right|_X && (2.8.36)
\end{aligned}$$

Here, a Maxwell's relation has been used (we can also start with expansion of the Jacobian). That is,

$$\left.\frac{\partial x}{\partial X}\right|_y < \left.\frac{\partial x}{\partial X}\right|_Y. \tag{2.8.37}$$

This implies that the indirect change occurs in the direction to reduce the effect of the perturbation (*Le Chatelier-Braun's principle*).[83]**q** A typical example is $C_V \leq C_P$:**q** larger specific heat implies that it is harder to warm up, that is, the system becomes more stable against heat injection. More generally, we may summarize the Le Chatelier-Braun principle as follows:

$$\begin{aligned}
(\Delta x)_y &\leq (\Delta x)_Y, && (2.8.38)\\
(\Delta X)_Y &\leq (\Delta X)_y. && (2.8.39)
\end{aligned}$$

The second formula is obtained by taking the reciprocal of (2.8.37).

To conclude the section on fluctuations, let us discuss the particle number fluctuation in dilute gas (or in a noninteracting particle system). It is a good occasion to discuss the law of small numbers. To study the exchange of particles between a 'small volume' and its surrounding the grand canonical partition function is convenient. The probability to find $n$ particles in the volume is

$$P(n) = \frac{1}{\Xi} Z_n e^{\beta n \mu}, \tag{2.8.40}$$

where $Z_n$ is the canonical partition function for $n$ particles that has the form $Z_n = (1/n!)f^n$, and

$$\Xi = \sum_n Z_n e^{\beta n \mu} = \exp\left(f e^{\beta\mu}\right). \tag{2.8.41}$$

---

[83] Karl Ferdinand Braun (1850-1918) [the inventor of the cathode-ray tube, the discoverer of principle of semiconductor diode, shared the Nobel prize with Marconi for wireless technology]. The history of this principle can be found in J. de Heer, "The principle of le Chatelier and Braun," J. Chem. Educ., **34**, 375 (1957). The form stated here is due to Ehrenfest.

Since$_{\mathbf{r}}$

$$\langle n\rangle = \frac{\partial \log \Xi}{\partial \beta\mu} = fe^{\beta\mu}, \tag{2.8.42}$$

we can rewrite $P(n)$ as

$$P(n) = \frac{1}{n!}\langle n\rangle^n e^{-\langle n\rangle}. \tag{2.8.43}$$

This is the *Poisson distribution* with average $\langle n\rangle$, which is also the variance of the distribution.

The same problem may be seen as follows. Let $\chi_i$ be the indicator of the event that $i$-th particle shows up in the volume we are looking at. Then,

$$n = \sum_{i=1}^{N} \chi_i \tag{2.8.44}$$

Here, $N$ is the number of particles in the 'world.' For any particle the probability for it to be in the volume is $\langle n\rangle/N$, so

$$P(n) = \binom{N}{n} p^n (1-p)^{N-n}. \tag{2.8.45}$$

For $N \gg n$ $\binom{N}{n} \simeq N^n/n!$, so in the $N \to \infty$ limit

$$P(n) = \frac{N^n}{n!}\left(\frac{\langle n\rangle}{N}\right)^n \left(1 - \frac{\langle n\rangle}{N}\right)^N \to \frac{1}{n!}\langle n\rangle^n e^{-\langle n\rangle}. \tag{2.8.46}$$

Mathematically, the Poisson distribution may be obtained for the observable $X$ whenever the following three conditions are satisfied:
(1) $X$ has an expectation value.
(2) $X$ is expressed as a sum of numerous independent random variables taking only the values 0 and 1.[84]
(3) None of these random variables can dominate the observable.[85]
This is called the *law of small numbers.*

---

[84] The distributions need not be the same.

[85] This means that, for example, '$x_1 \simeq \sum$ the rest' never occurs. Democracy is required for each realization (not only in law).

## 2.9 Chemical Reactions

Here, an elementary exposition of equilibrium chemical reactions is given.[r] Without chemical reactions no atomism was possible. Furthermore, the idea of detailed balance originated from chemical reactions. Also to understand chemical reactions is becoming increasingly important even for physicists because we living organisms are chemical machines.

Before discussing chemical reactions let us review chemical thermodynamics. The summary of thermodynamic laws in Section 1.1 deliberately confined itself to closed systems that do not have any exchange of matter with their environments. Now, we consider an *open system* that is allowed to exchange work, heat and matter with its environment. The zeroth and the fourth laws require no essential change. To describe an equilibrium state we need not only energy $E$ and work coordinates $\{X_i\}$ but also some extensive variables that specify the amount of chemical species $\{N_i\}$. Here, we simply assume that there is a set of independent variables $\{N_i\}$ that can uniquely specify the chemical composition of the system in equilibrium.

If the work coordinates and the variables $\{N_i\}$ can be treated exactly the same way, except for enlarging the set of independent extensive variables, almost no special care is needed. Simply we can define heat as the deficit of energy balance just as we reviewed in Section 1.1. Under the adiabatic condition if $N_i$ is changed by $dN_i$,[86] the internal energy of the system changes by $\mu_i dN_i$ (no summation convention). However, it is hard to construct an adiabatic wall that allows transferring chemicals across it. Therefore, let us consider the Helmholtz free energy change of a thermostatted system + thermostatted chemostats. In this case using the relation $W = \Delta A$ we can in principle measure the change of the free energy. Furthermore, we can change a particular $N_i$ infinitesimally. This requires a selective wall (*semipermeable wall*) that allows only a particular chemical to be exchanged across it. The existence of such a wall is assumed; at least approximately such a wall or a membrane exists. We could count many selective (but not energy-consuming) channels in biological systems as examples. Thus, the free energy change may be written as

$$dA = \sum_i x_i dX_i + \sum_i \mu_i dN_i, \tag{2.9.1}$$

where the second term is called the *mass action*.

---

[86] $N_i$ is a discrete variable, strictly speaking, but 'being macroscopic' implies that we cannot detect any discrete atomistic structure of matter empirically. Therefore, within thermodynamics, $N_i$ are never discrete variables.

**[T10]r The first law and the second law for an open system**[87] (**r** to ensembles, **r** to chemical reaction)

The first law may be extended as (here, thermodynamics for a closed system is assumed to have been formulated, so we can freely use thermodynamics without mass exchange or chemical reactions):

(i) The equilibrium state of a system is uniquely specified by a set of extensive variables including (internal) energy: $\{E, X_i, N_i\}$. Here, $X_i$ are = *work coordinates* and $N_i$ denotes the amount of chemical $i$.
(ii) There is a process called a reversible isothermal process such that $dA = \sum_i x_i dX_i + \sum_i \mu_i dN_i$.

The second law consists of the following two parts:

(i) When two equilibrium states are given, there is at least a process to change one state into the other through a combination (if necessary) of an isothermal process with mass action (to change the composition) and an adiabatic process without mass action (to change the temperature).
(ii) $W \geq \Delta A$ under the thermostatted condition, where $W$ is the sum of actual work and mass action,

Since the general formulas may be cumbersome, in this section, we use the following reaction to illustrate the general formulas:

$$N_2 + 3H_2 \leftrightarrow 2NH_3. \tag{2.9.2}$$

This formula implies that one molecule (or one mole[88]) of nitrogen reacts with 3 molecules (or 3 moles) of hydrogen to produce two molecules (or 2 moles) of ammonia. This does not mean that four molecules react at once; it is a summary of a certain number of elementary reactions.[89] The left hand side is called the *original system* (or *reactant system*) and the right hand side the *product system*. The coefficients 2, 3 and (not explicitly written) 1 (for nitrogen) are called *stoichiometric coefficients*. If we use the *sign convention* that the stoichiometric coefficients for the product system are all positive, and those for the original system all negative, we may write the reaction in an algebraic form

$$- \ N_2 - 3H_2 + 2NH_3 = 0. \tag{2.9.3}$$

---

[87] As stated before, reading the entries with [T] can provide an elementary (but not introductory) summary of thermodynamics.

[88] The quantity of one mole (mole number) is defined by the same number (Avogadro's number) of molecules as contained in 12 g of $^{12}$C).

[89] As the actual *elementary reactions* in the gas phase, such a reaction as (2.9.2) is very unusual, because elementary reactions are unimolecular decay or binary collision type reactions. The phenomenological scheme such as (2.9.2) may be interpreted as a time-coarse-grained reaction.

Thus, generally any reaction may be written as

$$\sum \nu_i A_i = 0, \tag{2.9.4}$$

where $\nu_i$ are *signed stoichiometric coefficients* for chemical $A_i$; $\nu_i > 0$ (resp., $\nu_i < 0$) implies $i$ is a product (resp., a reactant).

To describe the progress of a reaction, an extensive variable $\xi$ called the *extent of reaction* (or progress variable) is introduced as

$$d\xi = dn_i/\nu_i, \tag{2.9.5}$$

where $n_i$ is the quantity of chemical species $i$ in moles, and $dn_i$ is its change due to the reaction under consideration.[90] That is, $\nu_i\xi$ describes the quantity of the production of chemical $A_i$ ($\nu_i\xi > 0$ implies the production of chemical $i$ in the usual sense, and $\nu_i\xi < 0$ implies its consumption).

Let $b_i$ be a thermodynamic extensive quantity $B$ carried by the chemical species $i$ (often called molar quantity, e.g., molar enthalpy). Precisely, the quantity should be understood as

$$\partial B/\partial n_i = b_i \tag{2.9.6}$$

under constant $T, P$ and $n_j$ ($j \neq i$). Under constant $T$ and $P$[91], $B = \sum_i n_i b_i$ holds thanks to Euler's theorem. The change $\Delta B$ of the total $B$ due to the reaction (2.9.4) may be written as

$$\Delta B = \left(\sum \nu_i b_i\right) \Delta\xi. \tag{2.9.7}$$

For example, if $h_i$ is the molar enthalpy of chemical $i$

$$\Delta H = \left(\sum \nu_i h_i\right) \Delta\xi \tag{2.9.8}$$

is the *reaction heat*. Notice that $\Delta H > 0$ means that the system enthalpy increases, that is, enthalpy is absorbed by the proceeding of the reaction: it is an *endothermic* or *endoergic* reaction. If $\Delta H < 0$, the reaction is *exothermic* or *exoergic*. Recall Le Chatelier's principle. What is the sign of $\partial\xi/\partial T$, if the reaction is exothermic?

The conjugate variable $-\boldsymbol{A}$ of the extent of chemical reaction is called the *chemical*

---

[90] If there are several simultaneous reactions, $dn_i$ should carry the symbol denoting each reaction as $d^{(r)}n_i$ to describe the change due to the reaction $r$, and $d\xi^r = d^{(r)}n_i/\nu_i^r$, where $\nu_i^r$ is the signed stoichiometric coefficient for chemical $i$ in reaction $r$.

[91] These conditions must respect the actual experimental condition, but often chemical reaction experiments are done under constant $T$ and $P$.

*affinity*.[92] In terms of chemical potentials, the affinity for a particular reaction may be expressed as

$$\boldsymbol{A} = -\sum_i \nu_i \mu_i, \tag{2.9.9}$$

where $\mu_i$ is the chemical potential of chemical $i$.[93] Thus, the Gibbs relation reads

$$dE = TdS - PdV - \boldsymbol{A}d\xi + \cdots. \tag{2.9.10}$$

Under constant $T$ and $P$ the suitable thermodynamic potential is Gibbs' free energy $G$. The equilibrium condition for the reaction (2.9.4) is given by $\boldsymbol{A} = 0$:

$$-\boldsymbol{A} = \frac{\partial G}{\partial \xi} = \sum_i \frac{\partial n_i}{\partial \xi}\frac{\partial G}{\partial n_i} = \sum_i \nu_i \frac{\partial G}{\partial n_i} = \sum_i \nu_i \mu_i = 0. \tag{2.9.11}$$

To utilize this condition, we need expressions of chemical potentials.

We have already learned that for ideal gases,

$$\mu = RT \log[n/n_Q(T)] = RT \log[P/k_B T n_Q]. \tag{2.9.12}$$

For a gas mixture $P$ is replaced by the partial pressure $P_i$ of chemical $i$: we may write

$$\mu_i = \mu_i^\circ + RT \log P_i. \tag{2.9.13}$$

Here, $\mu_i^\circ$ is the chemical potential for $P_i = 1$.[94] In solutions the chemical potential of a solute $i$ in a solution is written as

$$\mu_i = \mu_i^\circ(T, P) + RT \log a_i, \tag{2.9.14}$$

where $a_i$ is called the *activity* of chemical $i$, which is close to the molarity $x_i$ when the solution is dilute.

The equilibrium condition for the reaction (2.9.4) reads

$$0 = \sum_i \nu_i [\mu_i^\circ(T, P) + RT \log a_i]. \tag{2.9.15}$$

---

[92] The minus sign in the definition is a convention, but it implies, "the reaction proceeds easier when the affinity is larger," so it is a convention familiar to chemists.

[93] If we have several reactions going on in parallel, for reaction $r$ we introduce the corresponding affinity $\boldsymbol{A}^r$. $-\boldsymbol{A}d\xi$ in (2.9.10) is replaced with $-\sum_r \boldsymbol{A}^r d\xi^r$, where $r$ denotes a particular reaction.

[94] If the gas is not a mixture of ideal gases, then the partial pressure must be replaced with the fugacity. Such chemical thermodynamics details will not be discussed here. See, for example, J. Kirkwood and I. Oppenheim, *Chemical Thermodynamics* (McGraw-Hill, 1961).

Or, (assuming the constant $T$ $P$ condition, so $\Phi = G$)

$$\boldsymbol{A}^\circ = -\Delta G^\circ \equiv -\sum_i \nu_i \mu_i^\circ(T,P) = RT\left(\prod_i a_i^{\nu_i}\right). \tag{2.9.16}$$

The left hand side does not depend on the chemical composition of the system, so we introduce the *chemical equilibrium constant* $K(T,P)$ according to

$$K(T,P) = e^{\boldsymbol{A}^\circ/RT} = \frac{\cdots a_p^{\nu_p}\cdots}{\cdots a_r^{-\nu_r}\cdots}, \tag{2.9.17}$$

where the numerator have all the products, and the denominator all the reactants. (2.9.17) is called the *law of mass action*. Note that all the exponents in the above formula are positive. Large $K$ implies that the reaction favors the product system in equilibrium (the reaction shifts to the right, or the standard affinity of the reaciton is large). The equilibrium constant for the reaction (2.9.2) is given by

$$K(T,P) = \frac{[\mathrm{NH_3}]^2}{[\mathrm{N_2}][\mathrm{H_2}]^3}. \tag{2.9.18}$$

Here, $[X]$ is the partial pressure (fugacity) of chemical $X$ in the gas phase reaction or the molarity (activity) in the solution.

If we differentiate the equilibrium constant with respect to $T$, we can obtain the *heat of reaction*, that is, $\Delta H$ due to reaction. The Gibbs-Helmholtz relation (or its analog for the Gibbs free energy) tells us

$$\left.\frac{\partial \log K}{\partial T}\right|_P = \frac{\Delta H^\circ}{RT}, \tag{2.9.19}$$

where $\Delta H^\circ$ is the enthalpy change for the 'standard state.' This is called *van't Hoff's equation*. Similarly,

$$\left.\frac{\partial \log K}{\partial P}\right|_T = -\frac{\Delta V^\circ}{RT}, \tag{2.9.20}$$

where $\Delta V^\circ$ is the volume change due to reaction for the 'standard state'. In reactions the change $\Delta$ always implies (product system) − (original system).

(2.9.19) tells us that if the reaction is exothermic (exoergic, i.e., $\Delta H^\circ < 0$), then increasing the temperature shifts the reaction to reduce the heat generation (i.e., $K$ decreases and the reaction tends to go back from the product system to the original). This is an example of *Le Chatelier's principle* asserting that "the response to a perturbation is in the direction to reduce its effect." (2.9.20) is also its example. Needless to say, these are manifestations of the stability of our world.

Consider a simple reaction

$$A \underset{k_-}{\overset{k_+}{\rightleftarrows}} B. \tag{2.9.21}$$

Here, $k_\pm$ are kinetic coefficients: if $A$ and $B$ are dilute, we may write

$$\frac{d[A]}{dt} = k_-[B] - k_+[A] = -\frac{d[B]}{dt}. \tag{2.9.22}$$

Therefore, in equilibrium two directional reactions dynamically balance in detail, and

$$k_+/k_- = [B]_{eq}/[A]_{eq} = K(T,P) = e^{-\Delta G^\circ/RT}, \tag{2.9.23}$$

where $K$ is the equilibrium constant. That is, the equilibrium constant can be described by the ratio of chemical reaction rates. However, never forget that this is not a general relation. This is true only under the assumption that the rate may be described directly in terms of concentrations. In general, reaction rates are assumed to be the function of activities instead of concentrations. Furthermore, when (2.9.22) is used away from equilibrium (as usually the case), the activity $a_i$ of chemical $i$ in equilibrium may not be the right activity to describe the reaction rate.

The equilibrium constant may be computed with the aid of (2.9.17). The relevant statistical mechanics problem is to compute the chemical potential of the standard state. In practice, this statistical mechanical computation can yield meaningful answers only for gas phase reactions, virtually when we may assume the gases are ideal. We have computed the chemical potential of monatomic ideal gas without internal degrees of freedom as (1.5.22). Chemical reactions among molecules with internal degrees of freedom are discussed in the following. This is not understandable before reading Section 3.6, so the reader may skip the topic.

To obtain the partition function for general ideal gases we have only to multiply the partition function for the internal degrees of freedom $z_I$ (see Sect. 3.6) to the usual ideal gas partition function (i.e., the partition function of monatomic ideal gas without internal degrees of freedom). Therefore, the chemical potential (when the amount of the substance is measured in moles) reads

$$\mu = \mu^\circ(T) + RT\log P, \tag{2.9.24}$$

where $P$ is the partial pressure and

$$\mu^\circ = -RT \log \left[ z_I \left( \frac{2\pi m k_B T}{h^2} \right)^{3/2} k_B T \right] \tag{2.9.25}$$

Here, $m$ is the mass of the molecule. We have only to use (2.9.17) and (2.9.16), but we must use the common energy origin for the chemical potentials for all the substances. The usual formula for $z_I$ uses the ground state of the molecule as the energy origin. Therefore, in order to use the usual $z_I$ (calculated as in Sect. 3.6) we must pay attention to the energy origin shifts due to, e.g., the difference in vibrational zero point energies.[r] That is, we need the energy difference (per molar extent of reaction) $\Delta$: the energy of the product system when all the products' internal states are in the ground states relative to the energy of the original system when all the reactants' internal states are in their ground states. Consequently, $\boldsymbol{A}^\circ$ must be replaced with $\boldsymbol{A}^\circ - \Delta$, if we stick to the conventional definition of $z_I$ whose energy origin is the ground state of the molecule. Instead of writing down a general expression, let us write down $K(T,p)$ for the ammonia synthesis reaction (2.9.2). Let $\Delta$ be the reaction energy for (2.9.2). Then, we may write

$$\begin{aligned} K(T,P) &= \exp\left[\left(\mu^\circ_{N_2} + 3\mu^\circ_{H_2} - 2\mu^\circ_{NH_3} - \Delta\right)/RT\right], & (2.9.26) \\ &= \frac{\left(m_{NH_3}^{3/2} z_{NH_3}\right)^2}{\left(m_{N_2}^{3/2} z_{N_2}\right)\left(m_{H_2}^{3/2} z_{H_2}\right)^3} \left(\frac{2\pi k_B T}{h^2}\right)^{-3} (k_B T)^{-2} e^{-\Delta/RT}. & (2.9.27) \end{aligned}$$

Here, $z_X$ is the partition function for the internal degrees of freedom for chemical species $X$. $\Delta$ is given by

$$\Delta = 2E^0_{NH_3} - 3E^0_{H_2} - E^0_{N_2}, \tag{2.9.28}$$

where $E^0_X$ is the ground state energy of chemical species $X$. These will be given in a problem of Chapter 3.

## Exercises for Chapter 2

**2.1** [On derivation of canonical distribution]
A textbook of statistical mechanics has the following derivation of the canonical distribution:
"The distribution function must not depend on the properties of particular systems, but must be universal. That is, the probability $P(E_{\rm I})$ (resp., $P(E_{\rm II})$) for the system I (resp., II) to have energy $E_{\rm I}$ (resp., $E_{\rm II}$) and the probability $P(E)$ for the compound system of I and II to have energy $E = E_{\rm I} + E_{\rm II}$ must have the same functional form. This must be so, as long as we expect statistical mechanics holds universally. Therefore,

$$P(E_{\rm I})P(E_{\rm II}) = P(E_{\rm I} + E_{\rm II}). \tag{2.P.1}$$

For this equation to be valid, we can prove that $P(E)$ must have the following functional form:

$$P(E) = Ce^{-\beta E}, \tag{2.P.2}$$

where $C$ is a constant."

Is this argument correct?

**2.2** [Elementary problem about spin system]
Due to the ligand field the degeneracy of the $d$-orbitals of the chromium ion $\mathrm{Cr}^{3+}$ is lifted, and the spin Hamiltonian has the following form

$$H = D(S_z^2 - S(S+1)/2), \tag{2.P.3}$$

where $D > 0$ is a constant with $S = 3/2$ (the cation is in the term $^4\mathrm{F}_{3/2}$).
(0) Why can you apply statistical mechanics to this 'single' ion?
(1) Compute the occupation probability of each energy level at temperature $T$ (you may use the standard notation $\beta = 1/k_BT$).
(2) Calculate the entropy.
(3) At high temperatures approximately we have $C = k_B(T_0/T)^2$ with $T_0 = 0.18$ K. Determine $D$ in K.

**2.3** [Vapor pressure of silicon]
The chemical potential $\mu_s$ of the atom in a solid is essentially given by the binding energy $\Delta$ of atom in the solid: $\mu_s = -\Delta$. Obtain the formula for the equilibrium vapor pressure of solid, and estimate the vapor pressure at room temperature of silicon for which $\Delta = 3$ eV.

**2.4** [Specific heat]

Suppose that a (3D) classical mechanical system has the following Hamiltonian

$$H = \sum_{i=1}^{N} a_k |\boldsymbol{p}_k + \boldsymbol{c}_k|^s \tag{2.P.4}$$

where $a_k$ ($k = 1, \cdots, N$), $s$ are positive constants, and $\boldsymbol{c}_k$ are constant 3-vectors. Without any explicit calculation compute the specific heat.

**2.5** [Permanent dipole]
The potential energy of a permanent electric dipole $\boldsymbol{p}$ is $U = -\boldsymbol{p} \cdot \boldsymbol{E}$ in the electric field $\boldsymbol{E}$. Obtain the electric susceptibility of the system.

**2.6** [Internal degree of freedom]
There is a classical ideal gas of volume $V$ consisting of $N$ molecules whose internal degrees of freedom are expressed by a single (quantum-mechanical) harmonic oscillator with a frequency $\nu$. The frequency depends on the volume of the system as

$$\frac{d\log \nu}{d\log V} = \gamma. \tag{2.P.5}$$

(1) Obtain the pressure of the system.
(2) Obtain the constant pressure specific heat $C_P$.

**2.7** [Application of equipartition of energy][95]
The internal motion of some ring puckering molecules (e.g., cyclobutanone) can be described by the following Hamiltonian:

$$H = \frac{p^2}{2m} + ax^4, \tag{2.P.6}$$

where $m$ is the effective mass of the oscillator and $a$ is a positive constant. Obtain the constant volume specific heat of this gas around the room temperature. Do not forget the contribution of translational degrees of freedom, etc.

**2.8**. [Equipartition of energy for relativistic gas]
For a classical relativistic ideal gas show that the counterpart of the law of equipartition of kinetic energy reads

$$\left\langle \frac{1}{2} \frac{mv_x^2}{\sqrt{1 - v^2/c^2}} \right\rangle = \frac{1}{2} k_B T. \tag{2.P.7}$$

**2.9** [An equality about canonical ensemble]

---

[95] original credit: B. Widom

Let $\Phi$ be the total potential energy of classical system. Show

$$\langle \Delta\Phi \rangle = \beta \left\langle (\nabla\Phi)^2 \right\rangle . \tag{2.P.8}$$

Here, the Laplacian and the nabla are understood as operators in the 3N-space.

**2.10** [Density operator for free particles: perhaps an elementary QM review]
The canonical density operator is given by

$$\rho = \frac{1}{Z} e^{-\beta H}, \tag{2.P.9}$$

where $H$ is the system Hamiltonian and $Z$ is the canonical partition function. Let us consider a single particle confined in a 3D cube of edge length $L$. We wish to compute the position representation of the density operator $\langle x|\rho|x'\rangle$ ($x$ and $x'$ are 3D position vectors, and bras and kets are normalized).

Let $U(\beta) = e^{-\beta H}$ and $H = p^2/2m$. There are two ways to compute $\langle x|U(\beta)|x'\rangle$:
A.
(1) Show that

$$\frac{\partial}{\partial\beta}\langle x|U(\beta)|x'\rangle = \frac{\hbar^2}{2m}\Delta_x \langle x|U(\beta)|x'\rangle, \tag{2.P.10}$$

where $\Delta_x$ is the Laplacian with respect to the coordinates $x$.
(2) What is the initial condition (i.e., $\langle x|U(0)|x\rangle$)?
(3) Solve the equation in (1) with the correct initial condition. You may use a simple boundary condition assuming the volume is very large (and temperature is not too low).
(4) Compute $Z$, using the result in (3). You may use (3) to study the finite volume system as long as the temperature is not too low.

B.
We can directly compute $\langle x|U(\beta)|x'\rangle$ with the aid of the momentum representation of $U(\beta)$:

$$\langle p|U(\beta)|p'\rangle = e^{-\beta p^2/2m}\delta(p-p').$$

(5) We use

$$\langle x|U(\beta)|x'\rangle = \int d^3p\, d^3p' \langle x|p\rangle\langle p|U(\beta)|p'\rangle\langle p'|x'\rangle.$$

What is $\langle x|p\rangle$? You may assume the infinite volume normalization (i.e., the $\delta$-function normalization: $\langle p|p'\rangle = \delta(p-p')$).
(6) Perform the integral in (5).

**2.11** [Density operator for a spin system]

Let $\rho$ be the density operator of a single 1/2 quantum spin system whose Hamiltonian[96] is given by $H = -\gamma \boldsymbol{\sigma} \cdot \boldsymbol{B}$, where $\boldsymbol{\sigma}$ is $(\sigma_x, \sigma_y, \sigma_z)$ in terms of the Pauli spin operators.
(1) Obtain the matrix representation of $\rho$ with respect to the base that diagonalizes $\sigma_z$.
(2) Find the average of $\sigma_y$.
(3) Obtain the matrix representation of $\rho$ with respect to the base that diagonalizes $\sigma_x$.

**2.12** [Legendre vs Laplace]
Consider an ideal gas consisting of $N$ atoms under constant pressure $P$ and temperature $T$.
(1) What is the most convenient partition function and the thermodynamic potential? Compute the partition function. You may use the ideal gas canonical partition function.
(2) Obtain the enthalpy of the system.

**2.13** [Constant magnetic field ensemble]
The following situation is the same as **1.18**: $N$ lattice sites have spins of $S = 1$ (in the term ${}^3P$), and the spin Hamiltonian has the following form:

$$H = DS_z^2. \tag{2.P.11}$$

(1) Consider this as the constant magnetic field ($B$) ensemble (also constant $T$ is assumed), and construct the corresponding generalized canonical partition function. The magnetization is given by $M = \mu \sum S_{zi}$.
(2) Compute the magnetization as a function of the magnetic field and temperature.

**2.14** [Absorption of mixed ideal gas, or convenient partition function][97]
There is a gas mixture consisting of two distinct atomic species $A$ and $B$. The mixture is an ideal gas and the partial pressures of $X$ is $p_X$ ($X = A$ or $B$). The gas is in equilibrium with an adsorbing metal surface on which there are adsorption sites. Atom X adsorbed at the site is with energy $-E_X$ on the average relative to the one in the gas phase, where $X = A$ or $B$. Each surface site can accommodate at most one atom. [Hint: you must know how to calculate the chemical potentials of the atoms, knowing the molecular weights.]
(1) Write down the 'partition function' (use the most convenient one) for the single site.
(2) Obtain the average surface concentration $n_X$ ($X = A$ or $B$) of atoms $A$ and $B$.

---

[96] Precisely speaking, this is te interaction Hamiltonian of the system with the magnetic field.

[97] Original credit: UIUC Qual F95.

(3) Under given (partial) pressures of $A$ and $B$ $n_E : n_A : n_B = 1 : 1 : 18$ (here, $E$ means empty). Find the maximum concentration $n_A$ obtainable with changing only the partial pressure of B.

**2.15** [Adsorption on catalytic surface]
There are $N$ adsorption centers on the catalyst surface exposed to a gas (ideal gas) of a certain chemical. Each adsorption center can accommodate at most two particles. The partition function for the single particle adsorption state is $a_1$ and the two particle adsorption state is $a_2$.
(1) Write down the single site (grand) partition function.
(2) Let $a_1 = 0$ (i.e., adsorption is possible only when a pair is formed). The average number of particles absorbed on the catalytic surface is $n_0$. Find the chemical potential of the particles.
(3) Now, the pressure of the chemical is doubled (with the temperature kept constant) and the average number of particles adsorbed on the catalytic surface is $n_1$. Find $n_1$ in terms of $N$ and $n_0$. $a_1$ is still assumed to be zero.
(4) If $a_1 > 0$, does the number of absorbed molecules increase from $n_0$ in (2) (i.e., the $a_1 = 0$ case)? Demonstrate your answer and give a brief physical explanation.

**2.16** [Gas under a weight]
Suppose there is a vertical cylindrical container of cross section $s$ whose top wall is a movable piston of cross section $s$ with mass $M$. The piston is assumed to move only in the vertical direction ($z$-direction) and feels gravity. The container contains $N$ ($\gg 1$) classical noninteracting particles with mass $m$.
(1) Write down the Hamiltonian of the gas + piston system (write the piston vertical momentum as $p_M$).
(2) Obtain the pressure $P$ of the gas, and write the Hamiltonian in terms of $P$ and the volume of the gas $V = sz$, where $z$ is the position of the piston from the bottom of the container.
(3) Now, the mechanical variables are the phase variables of the gas system and the piston momentum $p_M$ and $z = V/s$. Compute the canonical partition function of the whole system.
(4) You should have realized that the calculation in (3), apart from the unimportant contribution in the thermodynamic limit of the piston momentum, is the calculation of the pressure ensemble. [That is, the heavy piston acts as a constant pressure device.] Obtain the equation of state of the gas in the cylinder (a trivial question).

**2.17** [Ideal gas with the aid of grand canonical ensemble]
Let us study the classical ideal gas with the aid of the grand canonical ensemble. Let $\mu$ be the chemical potential.
(1) Compute the grand canonical partition function for a monatomic ideal gas. As-

sume that the mass of the atom is $m$.
(2) Find the internal energy and the pressure as a function of chemical potential $\mu$.
(3) Suppose the expectation value of the number of particles is $N$. How is the chemical potential determined?
(4) Are the results obtained above (especially the results of (2)) consistent with what you already know?

**2.18** [To obtain the microcanonical partition function with the aid of Laplace inverse transformation]
Starting from

$$Z = \frac{V^N}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2}, \tag{2.P.12}$$

obtain the microcanonical partition function $w(E, V)$ (with the aid of Laplace inverse transformation).

**2.19** [Equivalence of canonical and grand canonical ensembles]**r**
Let us check the equivalence of grand canonical and canonical ensembles. That is, if we compute thermodynamic quantities in the thermodynamic limit, both give the same answers. Even experimentalists should look at this proof at least once in their lives.

The grand partition function $\Xi(T, \mu)$ and canonical partition function $Z(T, N)$ (the ground state energy is taken to be the origin of energy) are related as

$$\Xi(T, \mu) = \sum_{N=0}^{\infty} Z(T, N) e^{\beta\mu N}.$$

Let us assume that the system consists of $N$ (which is variable) particles in a box of volume $V$ and the total interaction potential $\Phi$ among particles is bounded from below by a number proportional to the number of particles $N$ in the system: $\Phi \geq -NB$, where $B$ is a (positive) constant. (The system Hamiltonian generally has the form of $H = K + \Phi$, where $K$ is the kinetic energy.)

Through answering the following almost trivial questions, we can demonstrate the ensemble equivalence (rigorously).
(1) Show that there is a constant $a$ such that

$$Z(T, N) \leq \left(\frac{aV}{N}\right)^N. \tag{2.P.13}$$

Actually, show (classically)

$$Z(T, N) \leq Z_0(T, N) e^{\beta N B},$$

where $Z_0$ is the canonical partition function for the ideal gas (e.g., (1.7.3)). This is just eq.(6.2.105) above
(2) Show that the infinite sum defining the grand partition function actually converges. The reader may use eq.(6.2.105) and $N! \sim (N/e)^N$ freely.
(3) Choose $N_0$ so that

$$\sum_{N=N_0}^{\infty} Z(T,N)e^{\beta\mu N} < 1.$$

Show that this $N_0$ may be chosen to be proportional to $V$ (that is, $N_0$ is at most extensive).
(4) Show the following almost trivial bounds:

$$\max_N Z(T,N)e^{\beta\mu N} \le \Xi(T,\mu) \le (N_0+1)\max_N Z(T,N)e^{\beta\mu N}.$$

(5) We are almost done, but to be explicit, show that $PV/Nk_BT$ obtained thermodynamically from the canonical partition function and that directly obtained from the grand partition function agree.

**2.20** [Legendre transformation in convex analysis]
(1) We know that $-S$ is a convex function of internal energy $E$. Using the general property of the Legendre transformation, show that Helmholtz free energy $A$ is convex upward as a function of $T$. You may assume any derivative you wish to compute exists.
(2) When a phase transition occurs, the curve of $S(E)$ has a linear part as a function of $E$ (that is, $E$ can change under constant $T = T_e$). Then, $A$ as a function has a cusp at $T = T_e$ (that is, all the states corresponding to the flat part is collapsed to a point, the one-to-one nature of the Legendre transformation can be lost, if there is a phase transition). To illustrate this point, let us consider the following toy function

$$f(x) = \begin{cases} 2\tanh(x+1)-2 & \text{for } x < -1, \\ 2x & \text{for } -1 \le x \le 1, \\ (x-1)^2+2x & \text{for } x > 1. \end{cases}$$

Sketch its Legendre transform $f^*(\alpha) = \sup_x[\alpha x - f(x)]$. [Do not try to compute the explicit formula.]

**2.21** [Information]
Suppose there are two fair dice. We assume that one dice is red and the other is green (that is, distinguishable). Let us record the numbers that are up in this order as $(n, m)$ $(n, m \in \{1, 2, \cdots, 6\})$.
(1) To know a particular pair of numbers $(a, b)$ unambiguously what information (in

bits) do you need?
(2) You are told that the sum $a + b$ is not less than 5. What is the information you gain from this message?
(3) Next, you are told, one of the dice shows the face less than 3. What is the information you gain? (You must know the info obtained from (2) already.)
(4) Now, you are told that actually, the one of the dice in (3) is the red one. What is the information carried by this message?
(5) Finally, you are told that face pair is actually $(2, 5)$. What is the information in this final statement?

**2.22** [Variational principle for free energy (classical case)[98]]
Let $H = H_0 + V$ be a system Hamiltonian.
(1) Show that

$$A \le A_0 + \langle V \rangle_0,$$

where $A$ is the free energy of the system with $H$ and $A_0$ that with $H_0$. $\langle\ \rangle_0$ is the average over the canonical distribution of the system with the Hamiltonian $H_0$. The inequality is (sometimes) called the Gibbs-Bogoliubov inequality.
(2) We can use the inequality to estimate $A$. If we can compute $A_0$ and $\langle V \rangle_0$ (that is, the free energy for the system with $H_0$ and the average with respect to this system), then we can estimate the upper bound of $A$. Its minimum may be a good approximation to $A$. This is the idea of the variational approximation. Let us study an anharmonic oscillator with the Hamiltonian

$$H = \frac{1}{2m}p^2 + \frac{1}{2}kx^2 + \frac{1}{4}\alpha x^4,$$

where $m$, $k$ and $\alpha$ are positive constants. Let us define

$$H_0 = \frac{1}{2m}p^2 + \frac{1}{2}Kx^2.$$

Choose $K$ to obtain the best estimate of $A$ (you need not compute the estimate of $A$; it is easy but messy). You may use all the available results the text.

**2.23** [Gibbs-Bogoliubov's inequality (quantum case)][99]
Gibbs-Bogoliubov's inequality

$$A \le A_0 + \langle H - H_0 \rangle_0 \tag{2.P.14}$$

holds in quantum statistical mechanics as well.
(1) Demonstrate Peierls' inequality:

---

[98] This holds quantum mechanically as well, but the proof is not this simple.

[99] M D Girardeau and R M Mazo, "Variational methods in statistical mechanics,"

$$Tre^{-\beta H} \geq \sum_i e^{-\langle i|H|i\rangle}, \tag{2.P.15}$$

where $\{|i\rangle\}$ is an arbitrary orthonormal basis.
(2) Let $\{|i\rangle\}$ be the orthonormal basis consisting of the eigenstates of $H_0$. Then,

$$e^{-\beta A} \geq \sum_i e^{-\beta\langle i|H|i\rangle} = e^{-\beta A_0}\sum_i e^{\beta(A_0-\langle i|H_0|i\rangle)}e^{-\beta\langle i|(H-H_0)|i\rangle}. \tag{2.P.16}$$

Show Gibbs-Bogoliubov's inequality with the aid of Jensen's inequality.
**2.24** (1) For any density operator $\rho$

$$A \leq Tr[\rho(H + k_BT\log\rho)], \tag{2.P.17}$$

where $A$ is the free energy for the system whose hamiltonian is $H$.
(2) Suppose $\rho$ is the canonical density operator $\rho = e^{\beta(A_0-H_0)}$ for a system with the Hamiltonian $H_0$. Show that the above inequality is just Gibbs-Bogoliubov's inequality.

**2.25** [Convexity of free energy] (Ruelle)

$$A[\sum \lambda_i H_i] \geq \sum A[\lambda_i H_i]. \tag{2.P.18}$$

**2.26** [Thermodynamic perturbation theory]
Suppose the system Hamiltonian is given as $H = H_0 + \epsilon H_1$, where $\varepsilon$ is a (small) constant. Demonstrate the following expansion formula:

$$A = A_0 + \varepsilon\langle H_1\rangle_0 - \frac{1}{2}\beta\varepsilon^2\langle(H_1 - \langle H_1\rangle_0)^2\rangle_0 + \cdots, \tag{2.P.19}$$

where $A$ is the free energy of the system, $A_0$ is the free energy in case $H_1 = 0$, and $\langle\ \rangle_0$ is the expectation with respect to the canonical distribution with the Hamiltonian $H_0$.

**2.27** [Jarzynski's equality].[100]
A single stranded DNA with a certain binding protein is stretched slowly until the protein dissociates from the DNA. Then, the length of the DNA is returned slowly to the rather relaxed state where the binding of the molecule does not affect the

---

[100] Inspired by Rustem Khafizov and Yan Chemla's experiment on SSB. The numbers are only fictitious. although the magnitudes are realistic.

DNA tension. The work $W$ dissipated during the cycle is measured at 300 K and the experimental results were as follows:

| $W$ in pNnm | number of times | $\beta W$ | $e^{-\beta W}$ |
|---|---|---|---|
| 78-82 | 4 | 19.3 | $4.04 \times 10^{-9}$ |
| 83-87 | 15 | 20.5 | $1.21 \times 10^{-9}$ |
| 88-92 | 7 | 21.74 | $3.62 \times 10^{-10}$ |
| 93-97 | 4 | 22.94 | $1.082 \times 10^{-10}$ |
| 98-102 | 1 | 24.15 | $3.23 \times 10^{-11}$ |

What is the best estimate of the (Gibbs) free energy change due to binding of the protein in the relaxed state of the single stranded DNA? How is your estimate different from the simple average $\langle W \rangle$?

**2.28** [Fluctuation and spring constant][101]
Inside the $F_1$ATPase is a rotor $\gamma$ to which a long actin filament (it is a straight stiff bar of length 30 nm) is perpendicularly attached. Thus, the filament swings back and forth when the ATPase is waiting for an ATP molecule.
(1) The root mean square angle fluctuation of the stiff filament was 30 degrees at 290 K. If the temperature is raised by 10%, by what percentage will the angular fluctuation change? Assume that the molecular structure is not affected by this temperature change.
(2) What is the torsional spring constant of this rotor captured by the surrounding ring?
(3) Now, by adding appropriate polymers to the ambient solution, the effective viscosity of the solution is doubled. What is the mean square angle fluctuation of the filament? You may assume that the polymers do not affect the ATPase itself.

**2.29** [Thermodynamic fluctuations]
(1) Suppose $X$ and $y$ are nonconjugate pair with respect to energy, $X$ extensive and $y$ intensive. Prove that $\langle \delta X \delta y \rangle = 0$.
(2) Let $X$ and $x$ be a conjugate pair (wrt energy). Show $\langle \delta X \delta x \rangle = k_B T$.
(3) Express $\langle \delta \mu^2 \rangle$ in terms of a single thermodynamic derivative. The system is assumed to be described in terms of $S, V, N$ (or their conjugate variables).
(4) Show with the aid of grand partition function that

$$k_B T^2 \left.\frac{\partial E}{\partial T}\right|_{\mu,V} = \langle \delta E^2 \rangle - \mu \langle \delta E \delta N \rangle.$$

---

[101] If you wish to see the structure of the ATPase or to learn about how you can be alive, see K. Kinosita, K. Adachi, and H. Itoh, "Rotation of $F_1$ATPase: how an ATP-driven molecular machine may work," Ann. Rev. Biophys. Biomol. Struct., **33**, 245 (2005).

(5) Let $X$ be an extensive quantity. What can you conclude about $\langle \delta S \delta X \rangle$? The result is pedagogically suggestive, because entropy fluctuation means spatially local heat transport: that is, local temperature change.

**2.30** [Equilibrium fluctuation]
(1) Obtain $\langle \delta S \delta V \rangle$.
(2) Obtain $\langle \delta P \delta T \rangle$.

**2.31** [Fluctuation and Le Chatelier-Braun's principle]
(1) Show the following elementary algebraic inequality:

$$\langle \delta x \delta X \rangle^2 \le \langle \delta x^2 \rangle \langle \delta X^2 \rangle$$

where $x$ and $X$ make a conjugate pair of thermodynamic variables (with respect to energy).
(2) What is the relation between this inequality and the Le Chatelier-Braun principle? (Thus, you understand how natural the Le Chatelier-Braun principle is.)

**2.32** [Fluctuation of internal energy]
For a classical monatomic ideal gas consisting of $N$ atoms, compute the fluctuation of its internal energy (under constant $T$ and $P$). Or show

$$\langle (E - \langle E \rangle)^2 \rangle / \langle E \rangle^2 = 2/3N. \tag{2.P.20}$$

**2.33** [Stability and related topics, e.g., Le Chatelier-Braun]
(1) Suppose a phase transition from phase I to phase II occurs upon increasing the magnetic field in the $z$-direction. What can you say about the relation between the magnetisms of the phases?
(2) Suppose phase I is a low temperature phase and II a high temperature phase. The phase transition I → II is first order. What can you say about the sign of the latent heat $\Delta H$ of this phase transition?
(3) Which specific heat is larger, $C_B$ or $C_M$ (under constant magnetic field, and under constant magnetization, respectively)?
(4) Suppose there is a dielectric material between a parallel plate capacitor. The two plates of the capacitor may be short-circuited by a switch. What can you say about the relation between the heat capacity of the dielectric material under the open- and short-circuited conditions? Let $\varepsilon$ be its dielectric constant, that may or may not depend on temperature.
(5) Suppose there is a liquid that crystallizes upon heating. Discuss the latent heat

for this transition.[102]

**2.34** [Chemical equilibrium constant[103]]
The reaction

$$A \underset{k_-}{\overset{k_+}{\rightleftarrows}} B. \tag{2.P.21}$$

may be described as follows, if $A$ and $B$ are sufficiently dilute:

$$\frac{d[A]}{dt} = k_-[B] - k_+[A] = -\frac{d[B]}{dt}. \tag{2.P.22}$$

For all $t > 0$ show that

$$\frac{[B]_F(t)}{[A]_B(t)} = K, \tag{2.P.23}$$

holds. Here, $F$ denotes the forward reaction starting with pure A, and $R$ denotes the reverse reaction starting with the same moles of B as A. That is, if these two reactions are started simultaneously, then the concentration ratio at time $t$ as (6.2.158) is time-independent and equal to the chemical equilibrium constant. [However, this cannot be a general relation, but holds only under ideal solution and reaction conditions.]

---

[102] Johari, et al., "Endothermic freezing on heating and exothermic melting on cooling," J. Chem. Phys., **123**, 051104 (2005): $\alpha$-cyclodextrin + water + 4-methylpyridine (molar ratio of 1:6:100). For this system a liquid's endothermic freezing on heating and the resulting crystal's exothermic melting on cooling occur. $C_p$ decreases on freezing and increases on melting. Melting on cooling takes longer than freezing on heating.

[103] A. B. Adib, "Symmetry Relations in Chemical Kinetics Arising from Microscopic Reversibility," Phys. Rev. Lett., **96**, 028307 (2006).

# Chapter 3

# Non-interacting Systems

This chapter discusses the so-called ideal systems consisting of noninteracting particles.

That there are only two kinds of elementary particles, called *fermions* and *bosons*, is one of the fundamental principles of quantum mechanics. Let $P_{ij}$ be the particle exchange operator:

$$P_{ij}\psi(\cdots, x_i, \cdots, x_j, \cdots) = \psi(\cdots, x_j, \cdots, x_i, \cdots), \tag{3.0.1}$$

where $\psi$ is a many-body wave function, and $x_i$ is the position coordinates of particle $i$. Due to the indistinguishability of particles,[r] $P_{ij}$ does not change physics. Recall that not the vectors but the rays[1] determine physics. Therefore, we conclude

$$P_{ij}\psi(\cdots, x_i, \cdots, x_j, \cdots) = \lambda\psi(\cdots, x_i, \cdots, x_j, \cdots). \tag{3.0.2}$$

where $\lambda$ is a phase factor. $P_{ij}^2 = 1$ implies $\lambda^2 = 1$, so $\lambda = \pm 1$: if $+1$, the particles are bosons; if $-1$, fermions. This negative sign implies Pauli's exclusion principle: if there are two particles in the same state, the corresponding wave function of the

---

[1] ⟪**Rays**⟫ A ray is a one-dimensional subvector space spanned by a ket. In short, if two kets are identical up to the scalar multiplicative factors, they describe exactly the same physics except in interferences. That is, the phase factor does not matter. Interference between wave functions appear because we divide the whole world into subsystems (i.e., actually inseparable states are decomposed into two or more). Here, $\psi$ is the wave function for the system as a whole.

whole system must be identical to its negative version, so the function must be zero. That is, no such state is allowed. Fermions avoid each other in space, because they avoid occupation of the same one particle state. If the spin is a half odd integer, then the particle is a fermion; otherwise, it is a boson. This is called the *spin-statistics relation*.

The statistics of a composite particle (say, nuclei), if regarded as a single particle, are determined by its total spin. Therefore, if the particle consists of bosons and an odd number of fermions, the particle behaves as a single fermion; otherwise, as a single boson. Ejecting or absorbing particles, composite particles can switch their statistics.

We know that if we wish to understand phenomena with the scales between our length scale and the atomic scale, we have only to take into account the Coulomb interaction as interparticle interactions. For a charged particle system to be stable at least one of the charged species (say, the plus species) must be fermions. If both are bosons, the total energy of the system is not bounded from below (the system collapses).[2]

## 3.1 Grand canonical approach

It is convenient to discuss the ideal system with grand canonical ensemble.[3] Let us summarize the grand canonical formalism. We allow energy and particle exchange, so the appropriate Massieu function for a system with a constant volume is

$$S - \frac{E}{T} + \frac{\mu N}{T} = \frac{PV}{T}, \tag{3.1.1}$$

which is sometimes called Kramers' $q$. The corresponding generalized canonical partition function is the grand partition function:

$$\Xi = \sum_N \sum_E w_N(E) e^{-\beta(E-\mu N)}. \tag{3.1.2}$$

---

[2] **⟪Stability of matter⟫** E. H. Lieb, "The stability of matter," Rev. Mod. Phys. **48**, 553-569 (1976); "The Stability of Matter: From Atoms to Stars, 1989 Gibbs Lecture," Bull. Amer. Math. Soc., **22**, 1-49 (1990).

[3] Thanks to the ensemble equivalence, in the thermodynamic limit, microcanonical, canonical and grand canonical ensembles all give the same thermodynamics results, so this choice is a matter of convenience.

A convenient relation is

$$d\left(\frac{PV}{k_BT}\right) = d\log\Xi = -Ed\beta + \beta PdV + Nd(\beta\mu). \tag{3.1.3}$$

Let us consider the system consisting of noninteracting particles whose possible one-particle energies are in the set $\{\varepsilon_i\}$ (the degeneracy of the one particle states is taken into account by multiply placing the same energies in this list; $\varepsilon_i$ is the energy of the $i$-th state of a single particle). Since particles are not interacting, a microstate of our system may be completely described by the list of the *occupation numbers* $\{n_i\}$ of one particle states, where $n_i$ is the number of particles in the $i$-th one particle state. In the following do not confuse one particle quantum states and microstates of a macroscopic system.

The grand partition function may be computed as

$$\Xi(\beta,\mu) = \sum_N Z(N)e^{\beta\mu N}, \tag{3.1.4}$$

where the canonical partition function for the $N$-particle system reads[4]

$$Z(N) = \sum_E \left[\sum_{\sum n_i\varepsilon_i = E, \sum n_i = N} \prod_i e^{-\beta n_i\varepsilon_i}\right] = \sum_{\sum n_i = N} \prod_i e^{-\beta n_i\varepsilon_i}. \tag{3.1.5}$$

Here, the equations such as $\sum n_i\varepsilon_i = E$ describe constraints on $n_i$, when the sums are taken (these constraints are with some latitudes; for example $E$ has a latitude $\delta E$ as we discussed in the case of the microcanonical ensemble). Thus, we can write the grand partition function as

$$\Xi(\beta,\mu) = \sum_N \sum_{n_1+\cdots=N} \prod_i e^{-\beta(\varepsilon_i-\mu)n_i}, \tag{3.1.6}$$

$$= \prod_i \left(\sum_{n_i} e^{-\beta(\varepsilon_i-\mu)n_i}\right). \tag{3.1.7}$$

For bosons, the sum is over all nonnegative integers $n_i$, so

$$\Xi_{BE}(\beta,\mu) = \prod_i \left(1 - e^{-\beta(\varepsilon_i-\mu)}\right)^{-1}. \tag{3.1.8}$$

---

[4] Here, the degeneracy of energy levels are taken into account by counting the same level as many times as needed. Therefore, formally, the formulas with and without level degeneracies read the same.

Here, $BE$ implies ‘Bose-Einstein.’

On the other hand, for fermions, $n_i = 0$ or 1, so

$$\Xi_{FD}(\beta,\mu) = \prod_i \left(1 + e^{-\beta(\varepsilon_i - \mu)}\right). \tag{3.1.9}$$

Here, $FD$ implies ‘Fermi-Dirac.’

Thus, we can write as

$$\Xi(\beta,\mu) = \prod_i \left(1 \mp e^{-\beta(\varepsilon_i - \mu)}\right)^{\mp 1}. \tag{3.1.10}$$

Here, **the sign upstairs is for bosons, and the sign downstairs is for fermions**. We stick to this convention throughout this chapter.

Classic cases are small particle number density limits, so they correspond to the $\mu \to -\infty$ limit (cf. the Poisson distribution). Therefore, these grand partition functions both asymptotically go to

$$\Xi_{MB}(\beta,\mu) = 1 + \sum_i e^{-\beta(\varepsilon_i - \mu)}. \tag{3.1.11}$$

Here, $MB$ stands for ‘Maxwell-Boltzmann.’

The expectation value $\langle n_i \rangle$ is obtained from (3.1.10) as

$$\langle n_i \rangle = -\frac{\partial}{\partial(\beta\varepsilon_i)} \log \Xi = \frac{1}{e^{\beta(\varepsilon_i - \mu)} \mp 1}. \tag{3.1.12}$$

Take note the following points (here the lowest energy level is taken as the energy origin):

(i) For fermions, if the temperature is sufficiently low, $\langle n(\varepsilon) \rangle$ behaves like a step function $\theta(\mu - \varepsilon)$. Therefore, at sufficiently low temperatures chemical potential must be non-negative.

(ii) For bosons, the chemical potential cannot be positive, because $\langle n(\varepsilon) \rangle$ must be nonnegative. If the temperature is low enough, the one-particle ground state may be occupied by a macroscopic number of particles (the Bose-Einstein condensation as will be discussed in detail later).

(iii) The chemical potential decreases as temperature increases. Notice

$$\left.\frac{\partial \mu}{\partial T}\right|_N = -\left.\frac{\partial S}{\partial N}\right|_T. \tag{3.1.13}$$

Our system is ideal, so $S$ is proportional to $N$ and the entropy per particle is positive. Therefore, indeed $\mu$ decreases as $T$ increases. Thus, for fermions, $\mu(T)$ starts from some positive number and decreases as $T$ increases into the negative domain. For bosons, $\mu(T)$ starts from zero$_{\mathbf{q}}$ and decreases as $T$ increases.

The expected total number of particles $\langle N\rangle$ is obtained by differentiating the grand partition function wrt $\beta\mu$:

$$\langle N\rangle = \sum_i \frac{1}{e^{\beta(\varepsilon_i-\mu)} \mp 1}. \tag{3.1.14}$$

This is of course consistent with (3.1.12). Classically, $\mu \to -\infty$, so

$$\langle N\rangle = \sum_i e^{-\beta(\varepsilon_i-\mu)} = \log \Xi_{MB}, \tag{3.1.15}$$

but this is obvious from (3.1.11) and the equation of state of the classical ideal gas $PV = Nk_BT$.

The distinction between fermions and bosons show up clearly in pressure:

$$\frac{PV}{k_BT} = \log \Xi = \mp \sum_i \log\left(1 \mp e^{-\beta(\varepsilon_i-\mu)}\right). \tag{3.1.16}$$

(1) If $T, V, N$ are kept constant (the usually more interesting case than (2)), then

$$P_{FD} > P_{MB} > P_{BE}. \tag{3.1.17}$$

(2) If $T, V, \mu$ are kept constant, then

$$P_{BE} > P_{MB} > P_{FD}. \tag{3.1.18}$$

(2) is easy to see from $1/(1-x) > 1+x$ for $x \in (0,1)$ and (3.1.16). (1) requires some trick, so it will be shown below, but intuitively this can be understood by the extent of effective particle-particle attraction relative to the fermion case as the elementary two-box-two-particle system in Fig. 3.1.1 suggests. The figure not only suggests the pressures, it also suggests the extent of particle number fluctuations. As we will see later $\langle \delta n^2\rangle = \langle n\rangle(1-\sigma\langle n\rangle)$, where $\sigma$ is $+1$ for fermions, 0 for the classical limit and $-1$ for bosons.

(1) may be demonstrated as follows: Classically, $PV = Nk_BT$, so we wish to

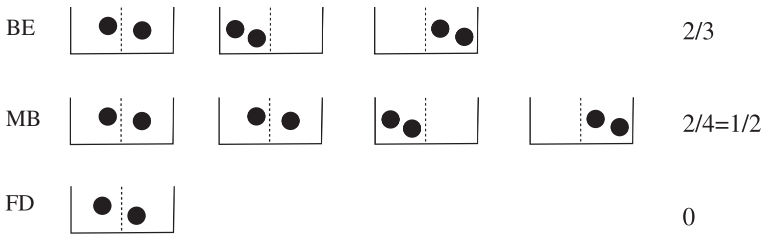


Fig. 3.1.1q Two-particle two box illustration of statistics. The numbers in the right denote the relative weights of the states for which effective attraction can be seen.

demonstrate ($N$ and $\langle N\rangle$ need not be distinguished, since we consider macrosystems)

$$\log \Xi_{FD} > \langle N\rangle > \log \Xi_{BE}. \tag{3.1.19}$$

Let us see the first inequality:[5] Writing $x_j = e^{-\beta(\varepsilon_j-\mu)}$, we have

$$\log \Xi_{FD} - \langle N\rangle = \sum_j \left[\log(1+x_j) - \frac{x_j}{1+x_j}\right], \tag{3.1.20}$$

where $x_j = e^{-\beta(\varepsilon_j-\mu)}$. We are done, because for $x>0$[6]

$$\log(1+x) - \frac{x}{1+x} > 0. \tag{3.1.21}$$

Similarly, we can prove the second inequality in (3.1.19).

(3.1.16) can be written in terms of the one-particle state density (translational degrees of freedom *only*) $D_t(\varepsilon)$ as

$$PV = \mp k_B T \int d\varepsilon\, D_t(\varepsilon) \log\left(1 \mp e^{-\beta(\varepsilon-\mu)}\right). \tag{3.1.22}$$

**Remark**. As we will learn later when the Bose-Einstein condensation occurs, the summation cannot always be replaced with integration, because the ground state contribution is significant. However, $\boldsymbol{p}=0$ particles should not contribute to $E$ nor to $P$ (as intuitively understandable), so the above formulas must be correct for very large volumes even if there is a Bose-Einstein condensation (so long as $\log N/N \ll 1$). We will come back to this point more quantitatively after the discussion of the Bose-Einstein condensation.

---

[5] The reader might wonder why we cannot use $\Xi_{MB}$ to demonstrate the formula; the reason is that $\mu$ in this grand partition function and that in $\Xi_{FD}$ or $\Xi_{BE}$ are distinct. Remember that we keep $N$; inevitably $\mu$ depends on statistics, so we cannot easily compare the Boltzmann factor $e^{\beta(\varepsilon-\mu)}$ in each term.

[6] Consider the derivatives.

A dimensional analysis may allow us to conclude $D_t(\varepsilon) \propto \varepsilon^{1/2}$ in 3-space, so

$$\int_0^\varepsilon D_t(\varepsilon) d\varepsilon = \frac{2}{3} D_t(\varepsilon)\varepsilon. \tag{3.1.23}$$

Using this in (3.1.22) with the aid of integration by parts (the boundary terms disappear), we can rewrite it as**rq**

$$PV = \frac{2}{3} E. \tag{3.1.24}$$

where $E$ is the internal energy

$$E = \sum_i \langle n_i \rangle \varepsilon_i = \int d\varepsilon\, D_t(\varepsilon) \frac{\varepsilon}{e^{\beta(\varepsilon-\mu)} \mp 1}. \tag{3.1.25}$$

We already know that (3.1.24) is true for the Maxwell-Boltzmann case, but now we know that it is true for any (non-relativistic) ideal gas (in 3-space). We will see the conclusion is a purely mechanical result with the aid of the virial theorem.

**Recalling $D_t(\varepsilon)$ quickly**
When $D_t(\varepsilon)$ is forgotten, perhaps the easiest method to recover it may be to use the relation between classical and quantum statistics:

$$\int_0^\varepsilon D_t(\varepsilon) d\varepsilon = \frac{1}{h^3} \int_{p^2/2m \le \varepsilon} d^3p d^3q = \frac{V}{h^3} \int_0^{p(\varepsilon)} 4\pi p^2 dp, \tag{3.1.26}$$

The integration with respect to the momenta is over 3D sphere of radium $\sqrt{2m\varepsilon}$

$$\int_0^\varepsilon D_t(\varepsilon) d\varepsilon = \frac{V}{h^3} \frac{4\pi}{3} (2m\varepsilon)^{3/2}. \tag{3.1.27}$$

That is,

$$D_t(\varepsilon) = 2\pi V \left(\frac{2m}{h^2}\right)^{3/2} \varepsilon^{1/2}. \tag{3.1.28}$$

**Remark**. In this book, the internal degrees of freedom of the particles, e.g., spins, are not taken into account in $D_t(\varepsilon)$. If there are $g$ states that are degenerate, then the answer must be multiplied by $g$. Since electrons are the most common examples, $g = 2$ due to spins is often built into $D_t(\varepsilon)$. However, in this book such $g$ will NEVER be included in $D_t$; we will write $gD_t$; we reserve $D_t$ for single particle translational aspects only. That is the reason for the subscript $t$.

## 3.2 Using annihilation and creation operators

We can go on without algebraic devices to compute various expectation values, but in any case, we must be familiar with the so-called second quantization formalism, so the needed device is introduced in the most efficient way.

Since we are studying indistinguishable particle systems, any microstate may be uniquely specified by the occupation numbers $n_k$ of one particle energy states $k$. That is, $\{n_k\}$ is a complete specification of a microstate of the system. Let us introduce a normalized ket $|n_k\rangle$ describing the state that the one-particle $k$-th energy level is occupied by $n_k$ particles. All the possible states allowed to the $k$-th single energy level is completely described by $n_k$ and the states with different occupation numbers can never be observed simultaneously, so we may assume that $\{|n_k\rangle\}_{n_k=0,1,\cdots}$ is an orthonormal complete set to describe this single energy level system. If we introduce a direct product ket $|\{n_k\}\rangle = |n_1\rangle|n_2\rangle\cdots$, we may assume that the totality of such kets makes a complete orthonormal system for the whole system. To describe the microstates choosing these kets as the basis is called *number representation*.

Consider a representative single one-particle energy level (the state name $k$ is omitted for simplicity), and write the state with $n$ particles on that level to be $|n\rangle$ without a suffix for simplicity. Let us *define*[7] a linear operator called the *creation operator* $a^+$ by

$$a^+|n\rangle = c_{n+1}|n+1\rangle. \tag{3.2.1}$$

Here, $c_{n+1}$ is a complex number we can choose conveniently. The adjoint of $a^+$ is denoted by $a$:

$$\langle n|a = c^*_{n+1}\langle n+1|. \tag{3.2.2}$$

(3.2.1) and (3.2.2) yield

$$\langle n|a|n+1\rangle = c^*_{n+1}\langle n+1|n+1\rangle = c^*_{n+1}. \tag{3.2.3}$$

For other combinations of bras and kets, this is zero. Therefore, $a|n+1\rangle$ must be parallel to $|n\rangle$: to be consistent with (3.2.3):

$$a|n+1\rangle = c^*_{n+1}|n\rangle. \tag{3.2.4}$$

---

[7] A linear operator on a certain vector space is defined if we specify all the matrix components with respect to a complete orthonormal set like $\{|n\rangle\}$. Thus, we can *freely* design an operator. Of course, whether it is useful or not is another matter.

This implies that $a$ may be called the *annihilation operator*. Since we cannot remove any particle from $|0\rangle$, we demand $c_0 = 0$:

$$a|0\rangle = 0. \tag{3.2.5}$$

(3.2.1) and (3.2.4) imply

$$a^+a|n\rangle = |c_n|^2|n\rangle. \tag{3.2.6}$$

**Boson case**

Now, we assume that $n$ is without upper bound, so $\{|n\rangle\}_{n\in\boldsymbol{N}}$ makes a complete orthonormal set. To define the operators we can freely choose $\{c_n\}$. A convenient choice is $c_n = \sqrt{n}$

$$a^+a|n\rangle = n|n\rangle. \tag{3.2.7}$$

With this choice $\hat{n} = a^+a$ is called the *number operator*. We see

$$aa^+ - a^+a = [a, a^+] = 1. \tag{3.2.8}$$

We can summarize the boson annihilation and creation operator as follows:

$$\begin{aligned} a^+|n\rangle &= \sqrt{n+1}|n+1\rangle, & (3.2.9)\\ a|n\rangle &= \sqrt{n}|n-1\rangle, & (3.2.10)\\ \hat{n} &= a^+a. & (3.2.11)\end{aligned}$$

The following commutation relations are useful:

$$[a, a^+] = 1,\ [a, a] = [a^+, a^+] = 0. \tag{3.2.12}$$

**Fermion case**

Pauli's principle tells us that $\{|0\rangle, |1\rangle\}$ is a complete set, so we may regard this as a complete orthonormal set. We may introduce a creation operator $a^+$ by

$$a^+|0\rangle = |1\rangle,\ \ a^+|1\rangle = 0. \tag{3.2.13}$$

The adjoint becomes an annihilation operator

$$a|0\rangle = 0,\ \ a|1\rangle = |0\rangle. \tag{3.2.14}$$

This defines $a$ and $a^+$ completely. We could write as

$$a^+|n\rangle = \sqrt{1-n}|n+1\rangle, \quad a|n\rangle = \sqrt{n}|n-1\rangle. \tag{3.2.15}$$

It is easy to see $\hat{n} = a^+a$ is the number operator. Since

$$\begin{aligned} (aa^+ + a^+a)|0\rangle &= |0\rangle, && (3.2.16) \\ (aa^+ + a^+a)|1\rangle &= |1\rangle, && (3.2.17) \end{aligned}$$

in terms of anticommutator $[A, B]_+ = AB + BA$

$$[a, a^+]_+ = 1, \ [a, a]_+ = [a^+, a^+]_+ = 0. \tag{3.2.18}$$

We must extend the above formalism to many one-particle energy state cases. Let $a_j$ and $a_j^+$ be the annihilation and creation operators for the $j$-th (one particle) energy state. To the end, we must specify the relation between $x_i x'_j$ and $x'_j x_i$, where $x$ and $x'$ are the creation or annihilation operators. The physical effect of these operations must be the same. We can freely choose the phase factor, so we 'uniformize' the rule:

Creation and annihilation operators: SUMMARY
**Bosons**:

$$\begin{aligned} a_i^+|n_1, \cdots, n_i, \cdots\rangle &= \sqrt{n_i + 1}|n_1, \cdots, n_i + 1, \cdots\rangle, && (3.2.19) \\ a_i|n_1, \cdots, n_i, \cdots\rangle &= \sqrt{n_i}|n_1, \cdots, n_i - 1, \cdots\rangle. && (3.2.20) \end{aligned}$$

The commutation relations are

$$[a_i, a_j] = [a_i^+, a_j^+] = 0, \quad [a_i, a_j^+] = \delta_{ij}. \tag{3.2.21}$$

$$|n_1, n_2, \cdots\rangle = \prod_{k=1} \frac{1}{\sqrt{n_k!}} (a_k^+)^{n_k} |0\rangle \tag{3.2.22}$$

makes a complete orthonormal set of a vector space called the *Fock space.*

**Fermions**:

$$\begin{aligned} a_i^+|n_1, \cdots, n_i, \cdots\rangle &= \sqrt{1 - n_i}|n_1, \cdots, n_i + 1, \cdots\rangle, && (3.2.23) \\ a_i|n_1, \cdots, n_i, \cdots\rangle &= \sqrt{n_i}|n_1, \cdots, n_i - 1, \cdots\rangle. && (3.2.24) \end{aligned}$$

These operators satisfy the following *anti-commutation* relations:

$$[a_i, a_j]_+ = [a_i^+, a_j^+]_+ = 0, \ \ [a_i, a_j^+]_+ = \delta_{ij}. \tag{3.2.25}$$

Analogous to (3.2.22), the space spanned by the orthonormal basis vectors

$$|n_1, n_2, \cdots\rangle = \prod_{k=1} (a_k^+)^{n_k} |0\rangle \tag{3.2.26}$$

is also called a *Fock space*.

Let $\psi_k(x) = \langle x|1_k\rangle$ be the (normalized) one particle wave function for the $k$-th state. Then,

$$\psi^+(x) = \sum_k \psi_k^*(x) a_k^+ \tag{3.2.27}$$

is an operator to create a particle at $x$. For bosons,

$$[\psi(x), \psi^+(y)] = \sum_k \psi_k(x)\psi_k^*(y) = \delta(x-y), \tag{3.2.28}$$

because $\{\psi_k\}$ is an orthonormal complete system (this is just Parseval's equality). All other commutators vanish. Needless to say, its fermion analogue holds with the anticommutator.

The Hamiltonian of the ideal system may be written as

$$H = \sum_r \varepsilon_r \hat{n}_r. \tag{3.2.29}$$

Since the total number of particles $N = \sum_r \hat{n}_r$, the grand partition function may be written as

$$\Xi = Tr e^{-\beta(H-\mu N)} = Tr \prod_r e^{-\beta(\varepsilon_r - \mu)\hat{n}_r}. \tag{3.2.30}$$

Here, $\hat{n}_r$'s commute, so we can factorize the product.

If we wish to compute $\langle \hat{n}_r\rangle$ we have only to compute

$$\langle \hat{n}_r\rangle = \frac{1}{\Xi_r} Tr\left[\hat{n}_r e^{-\beta(\varepsilon_r-\mu)\hat{n}_r}\right], \tag{3.2.31}$$

where $\Xi_r$ is the normalization constant. To compute this we use the following useful formulas (correct for both bosons and fermions) (the suffix $r$ is omitted for simplicity):

$$e^{A\hat{n}} a = a e^{A\hat{n}} e^{-A}, \tag{3.2.32}$$

$$e^{A\hat{n}} a^+ = a^+ e^{A\hat{n}} e^{A}. \tag{3.2.33}$$

Here, $A$ is anything commutative with the creation/annihilation operators. This implies

$$\langle aB\rangle = e^{\beta(\varepsilon-\mu)}\langle Ba\rangle, \tag{3.2.34}$$

$$\langle a^+B\rangle = e^{-\beta(\varepsilon-\mu)}\langle Ba^+\rangle. \tag{3.2.35}$$

Here, $B$ is an arbitrary operator. To show (3.2.34) we have only to compute

$$Tr\left[aBe^{-\beta(\varepsilon-\mu)\hat{n}}\right] = Tr\left[Be^{-\beta(\varepsilon-\mu)\hat{n}}a\right] = e^{\beta(\varepsilon-\mu)}Tr\left[Bae^{-\beta(\varepsilon-\mu)\hat{n}}\right] \tag{3.2.36}$$

Here, we have used the cyclic permutation invariance of $Tr(AB\cdots YZ) = Tr(ZAB\cdots Y)$.

For the boson case (If obvious the hat of the number operator may be dropped.)

$$1 = a_r a_r^+ - a_r^+ a_r \Rightarrow 1 = \langle a_r a_r^+\rangle - \langle n_r\rangle. \tag{3.2.37}$$

Thanks to (3.2.33)

$$\langle a_r a_r^+\rangle = e^{\beta(\varepsilon_r-\mu)}\langle n_r\rangle. \tag{3.2.38}$$

Therefore, (3.2.37) becomes

$$1 = e^{\beta(\varepsilon_r-\mu)}\langle n_r\rangle - \langle n_r\rangle. \tag{3.2.39}$$

That is,

$$\langle n_r\rangle = \frac{1}{e^{\beta(\varepsilon_r-\mu)} - 1}. \tag{3.2.40}$$

For fermions, the anti-commutation relation tells us

$$1 = a_r a_r^+ + a_r^+ a_r \Rightarrow 1 = \langle a_r a_r^+\rangle + \langle n_r\rangle. \tag{3.2.41}$$

Therefore,

$$\langle n_r\rangle = \frac{1}{e^{\beta(\varepsilon_r-\mu)} + 1}. \tag{3.2.42}$$

The expectation values of the products of creation and annihilation operators for noninteracting particle systems may be systematically reduced to the calculation of the average of number operators. The general formula is called the *Bloch-de Dominicis theorem.* In practice, the following iterative rule implied by the theorem is enough. Let $A_i$ be $a_i$ or $a_i^+$ (annihilation or creation operators). For fermions

$$\begin{aligned}\langle A_1A_2\cdots A_{2n}\rangle &= \langle A_1A_2\rangle\langle A_3\cdots A_{2n}\rangle - \langle A_1A_3\rangle\langle A_2A_4\cdots A_{2n}\rangle \\ &\quad + \langle A_1A_4\rangle\langle A_2A_3A_5\cdots A_{2n}\rangle + \cdots \pm \langle A_1A_{2n}\rangle\langle A_2\cdots A_{2n-1}\rangle.\end{aligned} \tag{3.2.43}$$

For bosons $-$-signs must be all $+$. For example, for the fermion case

$$\begin{aligned}\langle A_1A_2A_3A_4A_5A_6\rangle &= \langle A_1A_2\rangle\langle A_3A_4A_5A_6\rangle - \langle A_1A_3\rangle\langle A_2A_4A_5A_6\rangle + \langle A_1A_4\rangle\langle A_2A_3A_5A_6\rangle \\ &\quad -\langle A_1A_5\rangle\langle A_2A_3A_4A_6\rangle + \langle A_1A_6\rangle\langle A_2A_3A_4A_5\rangle\end{aligned} \tag{3.2.44}$$

and the 4-operator product can be reduced as

$$\langle A_1A_2A_3A_4\rangle = \langle A_1A_2\rangle\langle A_3A_4\rangle - \langle A_1A_3\rangle\langle A_2A_4\rangle + \langle A_1A_4\rangle\langle A_2A_3\rangle. \tag{3.2.45}$$

We can easily compute the variance of the occupation number (let us write $\langle\hat{n}\rangle = n$).**r** For bosons

$$\begin{aligned} \langle\hat{n}^2\rangle &= \langle a^+aa^+a\rangle = \langle a^+a\rangle\langle a^+a\rangle + \langle a^+a^+\rangle\langle aa\rangle + \langle a^+a\rangle\langle aa^+\rangle, && (3.2.46) \\ &= n^2 + n\langle aa^+\rangle = n^2 + n(1+n) = 2n^2 + n, && (3.2.47) \end{aligned}$$

so

$$\langle\delta n^2\rangle = n + n^2. \tag{3.2.48}$$

For fermions, we can use the same method, but since $\hat{n}^2 = \hat{n}$, we immediately obtain

$$\langle\delta n^2\rangle = n - n^2 = n(1-n). \tag{3.2.49}$$

For the classical case $n \ll 1$ for any one particle state, so

$$\langle\delta n^2\rangle = n. \tag{3.2.50}$$

This is in agreement with the result of the Poisson distribution as we already discussed in Chapter 1.

**The Bloch-de Dominicis theorem**

Here is the 'official statement' of the Bloch-de Dominicis theorem

Let $A_i$ be $a_i$ or $a_i^+$ (annihilation or creation operators for fermions). Then, we have for the canonical or grand canonical ensemble average of noninteracting (ideal) fermion systems (for boson systems sgn$(P)$ is simply replaced by unity everywhere; it is formally the same as *Wick's theorem* for Gaussian random variables.)

$$\langle A_1A_2\cdots A_{2n}\rangle = \sum_P \mathrm{sgn}(P)\langle A_{i_1}A_{i_2}\rangle\langle A_{i_3}A_{i_4}\rangle\cdots\langle A_{i_{2n-1}}A_{i_{2n}}\rangle, \tag{3.2.51}$$

where sgn$(P)$ is the sign of the permutation of $i_1, i_2, \cdots, i_{2n}$, and

$$\begin{aligned} i_1 < i_2, \quad i_3 < i_4 \quad \cdots, i_{2n-1} < i_{2n}, && (3.2.52) \\ i_1 < i_3 < \cdots i_{2n-1}. && (3.2.53) \end{aligned}$$

$\langle A_{i_1}A_{i_2}\rangle$ are called *contractions* and the sum in (3.2.51) is over all the ways to make contractions choosing pairs with the constraints above.

Here, we demonstrate the theorem for fermions. The case with bosons is analogous. We may assume

$$A_iA_j + A_jA_i = c(i,j), \tag{3.2.54}$$

where $c(i,j)$ is an ordinary number calculated by the anti-commutation relations.

Let $\langle\ \rangle$ be the canonical or grand canonical average. Then,

$$\begin{aligned}\langle A_1A_2\cdots A_{2n}\rangle &= \langle [c(1,2)-A_2A_1]\cdots A_{2n}\rangle, &(3.2.55)\\ &= c(1,2)\langle A_3\cdots A_{2n}\rangle - \langle A_2A_1A_3\cdots A_{2n}\rangle, &(3.2.56)\\ &= c(1,2)\langle A_3\cdots A_{2n}\rangle - \langle A_2[c(1,3)-A_3A_1]\cdots A_{2n}\rangle, &(3.2.57)\\ &= \cdots. &(3.2.58)\end{aligned}$$

Therefore, eventually, we obtain

$$\begin{aligned}\langle A_1A_2\cdots A_{2n}\rangle &= c(1,2)\langle A_3\cdots A_{2n}\rangle - c(1,3)\langle A_2A_4\cdots A_{2n}\rangle + \\ &\quad \cdots + c(1,2n)\langle A_2\cdots A_{2n-1}\rangle - \langle A_2A_3\cdots A_{2n}A_1\rangle. \quad (3.2.59)\end{aligned}$$

Now, let us assume that the average above is about the canonical or grand canonical ensemble and that the particles are non-interacting. That is, the system Hamiltonian is of the following form

$$H = \sum_r \varepsilon_r \hat{n}_r. \quad (3.2.60)$$

and (in the following for the grand canonical ensemble replace $\varepsilon_r$ with $\varepsilon_r - \mu$)

$$\langle A_2A_3\cdots A_{2n}A_1\rangle = \frac{1}{Z}Tr(A_2A_3\cdots A_{2n}A_1e^{-\beta H}). \quad (3.2.61)$$

Let $A_1 = a_1^{\pm}$, where $a_1^- = a_1$. Then,

$$a_1^{\pm}e^{-\beta H} = e^{-\beta H}e^{\pm\beta\varepsilon_1}a_1^{\pm}. \quad (3.2.62)$$

Therefore, using the cyclic permutation invariance of the trace operation, we have

$$\langle A_2A_3\cdots A_{2n}a_1^{\pm}\rangle = e^{\pm\beta\varepsilon_1}\langle a_1^{\pm}A_2A_3\cdots A_{2n}\rangle. \quad (3.2.63)$$

Therefore, (3.2.59) reads

$$\begin{aligned}(1+e^{\pm\beta\varepsilon_1})\langle A_1A_2\cdots A_{2n}\rangle &= c(1,2)\langle A_3\cdots A_{2n}\rangle - c(1,3)\langle A_2A_4\cdots A_{2n}\rangle + \\ &\quad \cdots + c(1,2n)\langle A_2\cdots A_{2n-1}\rangle. \quad (3.2.64)\end{aligned}$$

Now, consider

$$c(1,2)/(1+e^{\pm\beta\varepsilon_1}). \quad (3.2.65)$$

If $A_1 = a_1^+$, then $c(1,2)\neq 0$ only when $A_2 = a_1$. $\langle A_1A_2\rangle \neq 0$ only when $A_2 = a_1$. If $A_2 = a_1$, then $c(1,2) = 1$ and

$$\langle A_1A_2\rangle = \frac{1}{e^{\beta\varepsilon_1}+1} = \frac{c(1,2)}{e^{\beta\varepsilon_1}+1}. \quad (3.2.66)$$

If $A_1 = a_1$, then $c(1,2)\neq 0$ only when $A_2 = a_1^+$. $\langle A_1A_2\rangle \neq 0$ only when $A_2 = a_1^+$. If $A_2 = a_1^+$, then $c(1,2) = 1$ and

$$\langle A_1A_2\rangle = 1 - \frac{1}{e^{\beta\varepsilon_1}+1} = \frac{c(1,2)}{e^{-\beta\varepsilon_1}+1}. \quad (3.2.67)$$

Therefore,

$$c(1,2)/(1+e^{\pm\beta\varepsilon_1}) = \langle a_1^{\pm} A_2\rangle \tag{3.2.68}$$

for all cases. Thus, in (3.2.59) we may formally ignore the last term and replace all $c(1,j)$ with $\langle A_1 A_j\rangle$. It is obvious that we can repeat the same argument for smaller $n$, so the proof is essentially over.

Maxwell derived his distribution function based on the detailed balance argument.**q** Its essence is: if two particles in states $r$ and $r'$ collide and make two particles in states $s$ and $s'$, there must also be the reversed process as frequently as the original process (thanks to time reversal symmetry of mechanics). The collision process may be described by the interaction Hamiltonian $H_I$

$$H_I = \sum A_{s's,rr'} a_{s'}^+ a_s^+ a_r a_{r'}. \tag{3.2.69}$$

The representative term expresses a binary collision: $r$ and $r'$ states are destroyed and $s$ and $s'$ states are created. When particles satisfying the conditions collide, the probability amplitude for such a process actually happens is described by $A_{s's,rr'}$. If the initial ket is $|i\rangle = |\{n_k\}\rangle$, and if $r$, $r'$, $s$, and $s'$ are all distinct (i.e., the generic case),

$$\langle f|H_I|i\rangle = A_{s's,rr'}\sqrt{(n_s+1)(n_{s'}+1)n_r n_{r'}}. \tag{3.2.70}$$

Since $H_I$ is self-adjoint, it must also contain the term

$$A^*_{s's,rr'} a_{r'}^+ a_r^+ a_s a_{s'}. \tag{3.2.71}$$

This contributes

$$\langle f|A^*_{s's,rr'} a_{r'}^+ a_r^+ a_s a_{s'}|i\rangle = A^*_{s's,rr'}\sqrt{(n_{r'}+1)(n_r+1)n_s n_{s'}}. \tag{3.2.72}$$

In equilibrium if a transition A → B can occur, then the opposite transition B → A also can occur with the same probability (*detailed balance*). Therefore, (3.2.70) and (3.2.72) must have the same absolute value, or

$$(n_s+1)(n_{s'}+1)n_r n_{r'} = (n_{r'}+1)(n_r+1)n_s n_{s'}. \tag{3.2.73}$$

This implies

$$\frac{n_{s'}}{n_{s'}+1}\frac{n_s}{n_s+1} = \frac{n_{r'}}{n_{r'}+1}\frac{n_r}{n_r+1}. \tag{3.2.74}$$

That is,

$$q_s = \log\frac{n_s}{n_s+1} \tag{3.2.75}$$

is an additive invariant of the binary interaction (collision). Therefore, it is natural to assume that this is a linear function of energy:

$$\frac{n_s}{n_s+1} = e^{-\beta(\varepsilon-\mu)}. \tag{3.2.76}$$

This is exactly the Bose-Einstein distribution.

For fermions we have

$$\langle f|H_I|i\rangle = A_{s's,rr'}\sqrt{(1-n_s)(1-n_{s'})n_r n_{r'}}. \tag{3.2.77}$$

The detailed balance condition implies

$$(1-n_s)(1-n_{s'})n_r n_{r'} = (1-n_r)(1-n_{r'})n_s n_{s'}. \tag{3.2.78}$$

The same logic as before gives the Fermi-Dirac distribution.

The Maxwell's original case is $n_s n_{s'} = n_r n_{r'}$. This is consistent with the above quantum cases, because $n_s \ll 1$ for any state $s$ in the classical limit.

## 3.3 Ideal Fermion Systems

Let us summarize what we know.

The creation/annihilation operators:

$$[a_i, a_j]_+ = [a_i^+, a_j^+]_+ = 0,\ [a_i, a_j^+]_+ = \delta_{ij}, \tag{3.3.1}$$

and the particle number operator is

$$\hat{n}_j = a_j^+ a_j \tag{3.3.2}$$

with

$$f(\varepsilon) \equiv \langle \hat{n} \rangle = \frac{1}{e^{\beta(\varepsilon-\mu)}+1}, \tag{3.3.3}$$

where $\varepsilon$ is the energy of a particular one particle state (an eigenvalue of the single body Hamiltonian describing each particle) and $\mu$ is the chemical potential. The chemical potential (also called the *Fermi energy* or *Fermi level*) must be positive, if the temperature is sufficiently low, but is a monotone decreasing function of temperature.

The pressure is obtained with the aid of Kramers' $q$:

$$q = \frac{PV}{T} = k_B \sum_r \log(1 + e^{-\beta(\varepsilon_r - \mu)}). \tag{3.3.4}$$

Since $G = N\mu$,

$$A = G - PV = N\mu - k_B T \sum_r \log(1 + e^{-\beta(\varepsilon_r - \mu)}). \tag{3.3.5}$$

Also

$$\begin{aligned} N &= \sum_r \langle n_r \rangle = \sum_r f(\varepsilon_r), & (3.3.6)\\ E &= \sum_r \varepsilon_r \langle n_r \rangle = \sum_r \varepsilon_r f(\varepsilon_r). & (3.3.7) \end{aligned}$$

If we know the density of state, i.e., the density distribution function for the eigenvalues of the one particle Hamiltonian, $D_t(\varepsilon)$

$$\begin{aligned} A &= N\mu - k_B T \int d\varepsilon\, g D_t(\varepsilon) \log(1 + e^{-\beta(\varepsilon - \mu)}), & (3.3.8)\\ N &= \int d\varepsilon\, g D_t(\varepsilon) f(\varepsilon), & (3.3.9)\\ E &= \int d\varepsilon\, g D_t(\varepsilon) \varepsilon f(\varepsilon), & (3.3.10) \end{aligned}$$

where $g$ is the multiplicity due to internal degrees of freedom ($g = 2$ for electrons due to spin).

Intuitively, we may think that the total entropy is the sum of the contributions due to individual single particle states:

$$S = \sum_k S_k. \tag{3.3.11}$$

For each single state, $\langle n_k \rangle$ is the occupation probability, so we may guess with the aid of the Gibbs-Shannon formula:

$$S_k = -k_B[\langle n_k \rangle \log\langle n_k \rangle + (1 - \langle n_k \rangle) \log(1 - \langle n_k \rangle)]. \tag{3.3.12}$$

That is,

$$S = -k_B \int d\varepsilon\, gD_t(\varepsilon)[f(\varepsilon)\log f(\varepsilon) + (1 - f(\varepsilon))\log(1 - f(\varepsilon))]. \tag{3.3.13}$$

Our guess is correct. Let us derive the formula thermodynamically.

$$\begin{aligned} ST &= E - A = \int d\varepsilon\, gD_t(\varepsilon)\varepsilon f(\varepsilon) - \mu\int d\varepsilon\, gD_t(\varepsilon)f(\varepsilon) + k_BT\int d\varepsilon\, gD_t(\varepsilon)\log(1 + e^{-\beta(\varepsilon-\mu)}), && (3.3.14)\\ &= \int d\varepsilon\, gD_t(\varepsilon)(\varepsilon - \mu)f(\varepsilon) + k_BT\int d\varepsilon\, gD_t(\varepsilon)\log f(\varepsilon) - \int d\varepsilon\, gD_t(\varepsilon)(\varepsilon - \mu). && (3.3.15)\end{aligned}$$

Here, $k_BT\log(f(\varepsilon)^{-1} - 1) = \varepsilon - \mu$, so

$$ST = -k_BT\int d\varepsilon\, gD_t(\varepsilon)(1 - f(\varepsilon))\log\frac{1 - f(\varepsilon)}{f(\varepsilon)} - k_BT\int d\varepsilon\, gD_t(\varepsilon)\log f(\varepsilon). \tag{3.3.16}$$

Our guess is right.

Topics of some interest of the ideal fermion system may be:
(1) Qualitative feature of $f(\varepsilon)$ as a function of $T$.
(2) Properties at $T = 0$,
(3) Finite temperature corrections for small $T$.
The reader should be able to sketch (1) and (3).

Let us sketch $D_y(\varepsilon)f(\varepsilon)$ for various temperatures.

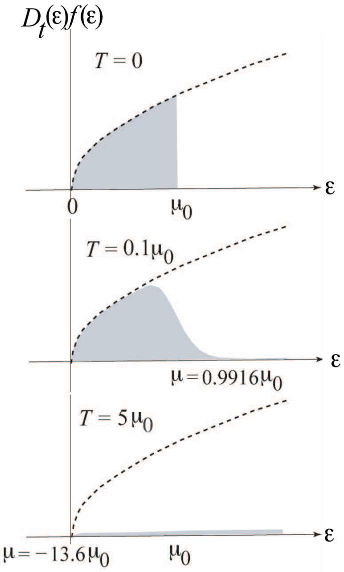


Fig. 3.3.1 $D_y(\varepsilon)f(\varepsilon)$
The dotted curve denotes the density of states $D_t$. The chemical potential decreases (toward $-\infty$) as $T$ increases. [The figure will be replaced with a better one later.]

The Fermi energy at $T = 0$ is implicitly defined by ($D_t$ is (3.1.28))

$$N = \int_0^{\mu(0)} gD_t(\varepsilon)f(\varepsilon) = V\frac{4}{3}\pi g\left(\frac{2m\mu(0)}{h^2}\right)^{3/2}. \tag{3.3.17}$$

The main feature is $n \propto \mu(0)^{3/2}$, which is equal to the number of one particle states below $\mu(0)$ and is proportional to the volume of the 3-ball of radium $\propto \sqrt{\mu(0)}$. (3.3.17) implies

$$\mu(0) = \frac{h^2}{2m}\left(\frac{3n}{4\pi g}\right)^{2/3}. \tag{3.3.18}$$

The density of states may be written as

$$gD_t(\varepsilon) = V\frac{3}{2}n\left(\frac{\varepsilon}{\mu(0)^3}\right)^{1/2}. \tag{3.3.19}$$

We know $D_t$ must be proportional to $\varepsilon^{1/2}$ and its dimension must be 1/energy, so $1/\mu(0)^{3/2}$ is required.

The internal energy at $T = 0$ must be $\propto n\mu(0)$ per volume.**q** Indeed,

$$e(0) = \frac{1}{V}\int_0^{\mu(0)} \varepsilon g D_t(\varepsilon)d\varepsilon = \frac{3}{2}n\mu(0)^{-3/2}\int_0^{\mu(0)} \varepsilon^{3/2}d\varepsilon = \frac{3}{5}\mu(0)n. \tag{3.3.20}$$

This can be derived quickly from the form $D_t = A\varepsilon^{1/2}$ ($A$ is a constant independent of $\varepsilon$); $N = (2/3)A\mu(0)^{3/2}$ and $E = (2/5)A\mu(0)^{5/2}$, so $E = (3/5)\mu(0)N$. $E/V$ immediately gives us the pressure at $T = 0$

$$PV = \frac{2}{5}\mu(0)N. \tag{3.3.21}$$

Now, we must compute the correction due to small $T$ for various quantities. The basic idea is that at low temperatures, the cliff of the Fermi distribution is steep, so its derivative is close to a $\delta$-function around $\mu$.

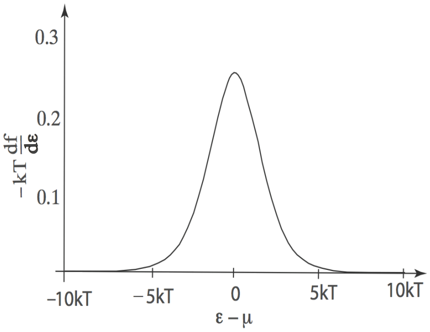


Fig. 3.3.2 The derivative of the Fermi distribution. Its width is about $5k_BT$. and the height is $\beta/4$.

Thus,

$$J = \int_0^\infty d\varepsilon\phi(\varepsilon)\left(-\frac{df}{d\varepsilon}\right) \simeq \phi(\mu). \tag{3.3.22}$$

To improve this approximation we Taylor expand $\phi$ around $\mu$ as

$$J = -\int_{-\infty}^{+\infty} d\varepsilon\left[\phi(\mu) + (\varepsilon-\mu)\phi'(\mu) + \frac{1}{2}(\varepsilon-\mu)^2\phi''(\mu) + \cdots\right]\frac{df}{d\varepsilon}. \tag{3.3.23}$$

Exchanging the order of integration and expansion, we have only to compute the integral of the power $(\varepsilon-\mu)^n$. $df/d\varepsilon$ is symmetric around $\mu$, so the odd powers vanish. The result is a useful formula for $J$:

$$-\int_{-\infty}^{+\infty} d\varepsilon\,\phi(\varepsilon)\frac{df(\varepsilon)}{d\varepsilon} = \phi(\mu) + (k_BT)^2\zeta(2)\phi^{(2)}(\mu) + \cdots. \tag{3.3.24}$$

Here, $\zeta$ is the zeta function; $\zeta(2) = \pi^2/6$.

Analogously, we should be able to expand the following integral, if $\phi(0) = 0$ and the asymptotic increase rate of $\phi$ is at most algebraic:q

$$\int_0^{+\infty} d\varepsilon\,\phi(\varepsilon)f(\varepsilon) = \int_0^\mu \phi(x)dx + (k_BT)^2\zeta(2)\phi'(\mu) + \cdots. \tag{3.3.25}$$

**The full expansion formula**

$$J = -\int_{-\infty}^{+\infty} d\varepsilon\, \phi(\varepsilon)\frac{df(\varepsilon)}{d\varepsilon} = \phi(\mu) + \sum_{r=1}^{\infty} 2(1-2^{1-2r})\zeta(2r)\phi^{(2r)}(\mu)(k_BT)^{2r} \quad (3.3.26)$$

may be obtained as follows. Taylor-expanding $\phi$, we obtain

$$J = \phi(\mu) - \sum_{n=1}^{\infty}\int_{-\infty}^{\infty} d\varepsilon\, \frac{1}{n!}(\varepsilon-\mu)^n\frac{df}{d\varepsilon}\phi^{(n)}(\mu) = \phi(\mu) + \sum_{n=1}^{\infty}(k_BT)^n\frac{1}{n!}K_n\phi^{(n)}(\mu). \quad (3.3.27)$$

Here,

$$K_n \equiv -\int_{-\infty}^{+\infty} d\varepsilon\,(\varepsilon-\mu)^n\frac{df}{d\varepsilon}. \quad (3.3.28)$$

This is zero for odd $n$. We have

$$K_n = -2(k_BT)^n\int_0^{\infty} dy\, y^n\frac{d}{dy}\frac{1}{e^y+1} = 2n(k_BT)^n\int_0^{\infty} dy\,\frac{y^{n-1}}{e^y+1}. \quad (3.3.29)$$

for even $n > 0$. To compute this we use the following expansion

$$\frac{1}{e^y+1} = \frac{e^{-y}}{1+e^{-y}} = \sum_{k=0}^{\infty}(-1)^k e^{-(k+1)y}, \quad (3.3.30)$$

and the integral and this expansion may be exchanged: we can use

$$\int_0^{\infty} dy\, y^{n-1}e^{-(k+1)y} = \frac{\Gamma(n)}{(k+1)^n}. \quad (3.3.31)$$

Therefore,

$$K_n = 2(k_BT)^n n!\sum_{k=0}^{\infty}(-1)^k\frac{1}{(k+1)^n}. \quad (3.3.32)$$

The sum can be computed as

$$\begin{aligned}\sum_{k=0}^{\infty}(-1)^k\frac{1}{(k+1)^n} &= \sum_{\text{even}}\frac{1}{(k+1)^n} - \sum_{\text{odd}}\frac{1}{(k+1)^n} = \sum_{\text{odd}}\frac{1}{k^n} - \sum_{\text{even}>0}\frac{1}{k^n} && (3.3.33)\\ &= \sum_{k=0}^{\infty}\frac{1}{(2k+1)^n} - \sum_{k=1}^{\infty}\frac{1}{(2k)^n} = \sum_{k=1}^{\infty}\frac{1}{k^n} - 2\sum_{k=1}^{\infty}\frac{1}{(2k)^n} && (3.3.34)\\ &= \zeta(n) - 2^{1-n}\zeta(n). && (3.3.35)\end{aligned}$$

Therefore,

$$K_n = 2(k_BT)^n n!(1-2^{1-n})\zeta(n). \quad (3.3.36)$$

We may apply (3.3.25) to $N$ because $D_t$ grows only algebraically ($\zeta(2) = \pi^2/6$):

$$N = \int_0^\infty d\varepsilon\, gD_t(\varepsilon)f(\varepsilon) = \int_0^\mu dx\, gD_t(x) + \zeta(2)\left.\frac{dgD_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu}(k_BT)^2 + \cdots. \tag{3.3.37}$$

The Fermi level $\mu(0)$ at $T = 0$ has already been computed with the aid of

$$N = \int_0^{\mu(0)} gD_t(\varepsilon)d\varepsilon. \tag{3.3.38}$$

Therefore, combining with (6.3.71), we can find $\mu$ as a function of $T$: First, we see

$$\int_0^{\mu(0)} gD_t(\varepsilon)d\varepsilon = \int_0^\mu dx\, gD_t(x) + \zeta(2)\left.\frac{dgD_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu}(k_BT)^2 + \cdots. \tag{3.3.39}$$

We can guess $\mu = \mu(0) + a(k_BT)^2 + \cdots$, so substituting this relation in the above formula to fix $a$, we obtain the low temperature formula for the Fermi level:

$$\mu(T) = \mu(0) - \frac{\pi^2}{6}\frac{d}{d\varepsilon}\log gD_t(\varepsilon)\Big|_{\varepsilon=\mu(0)}(k_BT)^2 + \cdots. \tag{3.3.40}$$

This may be streamlined as follows for sufficiently low temperature:

$$\mu(T) = \mu(0)\left[1 - \frac{\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right]. \tag{3.3.41}$$

Here, $T_F = \mu(0)/k_B$ is the *Fermi temperature*.[8] This behavior can be guessed from (i) no change of any thermodynamic potential (per particle) at $T = 0$ (hence, no first order derivative wrt $T$ is nonzero) and (ii) for $T > 0$ $\mu$ is a decreasing function of $T$.

(3.3.25) allows us to calculate temperature corrections such as

$$\begin{aligned} E &= \int d\varepsilon\, gD_t(\varepsilon)\varepsilon f(\varepsilon) && (3.3.42)\\ &= \int_0^\mu d\varepsilon gD_t(\varepsilon)\varepsilon + \zeta(2)(k_BT)^2\left.\frac{dgD_t(\varepsilon)\varepsilon}{d\varepsilon}\right|_\mu + \cdots. && (3.3.43)\end{aligned}$$

---

[8] Representative values: Li $5.5\times10^4$K (4.7 eV); Cu $8.2\times10^4$K (8.2 eV); Au $5.5\times10^4$K (5.5 eV).

Therefore, introducing (3.3.40) and expanding the integral around $\mu(0)$, we obtain (notice that $D_t'(\varepsilon)\varepsilon = D_t(\varepsilon)/2$)**q**

$$E = E_0 + \zeta(2) g D_t(\mu(0))(k_B T)^2 + \cdots, \tag{3.3.44}$$

where $E_0 = Ve(0)$ is the zero-point energy. A more streamlined low temperature formula for the internal energy density reads

$$e(T) = \frac{3}{5} n\mu(0) \left\{ 1 + \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \cdots \right\}. \tag{3.3.45}$$

Its high temperature asymptote reads

$$e(T) \simeq \frac{3}{2} n k_B T, \tag{3.3.46}$$

so we may sketch $E(T)$ as Fig. 3.3.3.

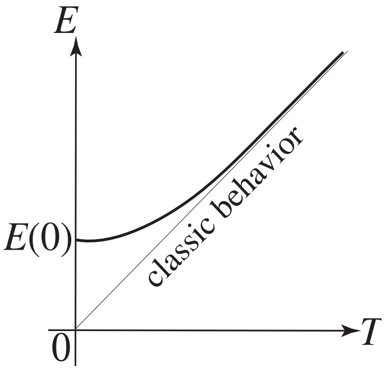


Fig. 3.3.3**q** Internal energy of ideal Fermi gas as a function of $T$.

From the internal energy obtained in (3.3.45), it is easy to see that the low temperature specific heat reads

$$C_V \simeq \frac{\pi^2}{3} k_B^2 T g D_t(\mu(0)). \tag{3.3.47}$$

This reads

$$C_V \simeq \frac{1}{2}\pi^2 n k_B^2 T/\mu(0) = \frac{1}{2}\pi^2 n k_B \frac{T}{T_F} \tag{3.3.48}$$

for ideal Fermi gas, where $T_F$ is the Fermi temperature; the dependence on spin is only through $T_F$. $C_V \sim T/T_F$ can easily be guessed dimensional analytically as well.

Specific heat is proportional to the number of thermally excitable degrees of freedom (in ideal gas this is proportional to the number of thermally excitable particles); the reader must be able to explain $\propto T$ intuitively. For ordinary metals, however, it is not easy to see this linear behavior, because $T_F$ is very high and easily masked by the phonon contribution.

It is wise to use the specific heat to compute the $T$ dependence of entropy:

$$S(T) = \int_0^T \frac{C_V}{T} dT. \tag{3.3.49}$$

Therefore, at low temperatures

$$S(T) \simeq C_V(T). \tag{3.3.50}$$

For sufficiently low temperatures, we can use the formula for internal energy (3.3.45) to obtain the equation of state.

$$PV = \frac{2}{5} N k_B T_F \left(1 + \frac{5}{12}\pi^2 \left(\frac{T}{T_F}\right)^2 + \cdots\right). \tag{3.3.51}$$

Clearly recognize that this is extremely large for metals.

For sufficiently high temperatures (or rather, for sufficiently low chemical potential close to classical limit) fugacity $z = e^{\beta\mu}$ is very small, so we wish to expand $PV$ and $N$, $T$ in powers of fugacity. $PV$ in terms of the fugacity $z$ is:

$$\frac{PV}{k_B T} = \int d\varepsilon\, g D_t(\varepsilon) \log(1 + z e^{-\beta\varepsilon}) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} z^n}{n} \int d\varepsilon e^{-n\beta\varepsilon} g D_t(\varepsilon), \tag{3.3.52}$$

and

$$N = \int d\varepsilon\, g D_t(\varepsilon) \frac{1}{1 + e^{\beta\varepsilon}/z} = \sum_{n=1}^{\infty} (-1)^{n-1} z^n \int d\varepsilon e^{-n\beta\varepsilon} g D_t(\varepsilon). \tag{3.3.53}$$

Both are analytic functions of $z$ along the positive real $z$-axis.[9] Solving $z$ in powers of number density $n$, we can obtain the *virial expansion* of the equation of state for the ideal Fermi gas:

$$\frac{PV}{NRT} = 1 + \frac{1}{2^{5/2} g} \frac{n}{n_Q} + \cdots. \tag{3.3.54}$$

Here, $g$ is the spin degeneracy, and $n_Q = (2\pi m k_B T/h^2)^{3/2}$ as before.

---

[9] Hence, there is no phase transition. See Chapter 5.

## 3.4 Ideal Boson Systems

The creation/annihilation operators (→Section 3.2):

$$[a_i, a_j] = [a_i^+, a_j^+] = 0,\ [a_i, a_j^*] = \delta_{ij}, \tag{3.4.1}$$

and the particle number operator is

$$\hat{n}_j = a_j^* a_j \tag{3.4.2}$$

with

$$\langle \hat{n}_j \rangle = \frac{1}{e^{\beta(\varepsilon_j - \mu)} - 1}. \tag{3.4.3}$$

The chemical potential $\mu$ is a monotone decreasing function of $T$ and cannot be larger than the one-particle ground-state energy, because $\langle n(\varepsilon) \rangle$ must be positive.

The pressure is obtained by

$$q = \frac{PV}{T} = -k_B \sum_r \log(1 - e^{\beta(\varepsilon_r - \mu)}). \tag{3.4.4}$$

Since $G = N\mu$,

$$A = G - PV = N\mu + k_B T \sum_r \log(1 - e^{-\beta(\varepsilon_r - \mu)}). \tag{3.4.5}$$

Also

$$N = \sum_r \langle n_r \rangle = \sum_r f(\varepsilon_r), \tag{3.4.6}$$

$$E = \sum_r \varepsilon_r \langle n_r \rangle = \sum_r \varepsilon_r f(\varepsilon_r), \tag{3.4.7}$$

where

$$f(\varepsilon) = \frac{1}{e^{\beta(\varepsilon - \mu)} - 1}. \tag{3.4.8}$$

$\mu$ must be a function of $N$ and $T$ for a canonical ensemble. We know

$$PV = \frac{2}{3} E \tag{3.4.9}$$

is true for any (non-relativistic) ideal gas in 3-space.

Intuitively, we may think that the total entropy is the sum of the contributions due to each one particle state:

$$S = \sum_k S_k, \tag{3.4.10}$$

where $S_k$ is the entropy contribution by a one particle energy state $k$ with energy $\varepsilon_k$. Let $f = \langle \hat{n}_k \rangle$. The Gibbs-Shannon formula tells us that

$$S_k/k_B = -\sum_n p_n \log p_n, \tag{3.4.11}$$

where $p_n$ is the probability for this energy state to have $n$ bosons. This can be computed as (suffixes $k$ are omitted for simplicity)

$$p_n = e^{-n\beta(\varepsilon-\mu)}/(1 + e^{-\beta(\varepsilon-\mu)} + e^{-2\beta(\varepsilon-\mu)} + \cdots) = \frac{1}{1+f}\left(\frac{f}{1+f}\right)^n. \tag{3.4.12}$$

Putting this into (3.4.11), we obtain

$$\begin{aligned} S_k/k_B &= -\sum_n \frac{1}{1+f}\left(\frac{f}{1+f}\right)^n \left[-\log(1+f) + n\log\frac{f}{1+f}\right] \\ &= \log(1+f) - f\log\frac{f}{1+f} = (1+f)\log(1+f) - f\log f. \end{aligned} \tag{3.4.13}$$

That is (compare the following result with (3.3.13)),

$$S = -k_B\sum_k [f(\varepsilon_k)\log f(\varepsilon_k) - (1+f(\varepsilon_k))\log(1+f(\varepsilon_k))]. \tag{3.4.14}$$

This indeed agrees with the thermodynamic result:

$$ST = E - A = \sum_k \varepsilon_k f(\varepsilon_k) - \mu\sum_k f(\varepsilon_k) - k_BT\sum_k \log(1-e^{-\beta(\varepsilon_k-\mu)}) \tag{3.4.15}$$

$$= -k_BT\sum_k\left\{f(\varepsilon_k)\log\frac{f(\varepsilon_k)}{1+f(\varepsilon_k)} + \log\frac{1}{1+f(\varepsilon_k)}\right\}. \tag{3.4.16}$$

We have already noted that rewriting the sum $\sum_k \to \int d\varepsilon D_t(\varepsilon)$ may not be correct for bosons. Let us explicitly see why. Let us formally rewrite the formula for $N$ as

$$N_1(T,\mu) = \frac{V}{h^3}\int_0^\infty \frac{1}{e^{\beta(p^2/2m-\mu)}-1}4\pi p^2 dp. \tag{3.4.17}$$

This integral is a monotone increasing function of $\mu$, so its upper bound can be evaluated as[q]

$$N_1(T,\mu) \le N_1(T,0) = \frac{V}{h^3}\int_0^\infty \frac{1}{e^{\beta p^2/2m}-1}4\pi p^2 dp = AVT^{3/2}, \tag{3.4.18}$$

where $A$ is a positive constant. Thus, if $T$ is sufficiently low $N_1 < N$. That is, the integral approximation cannot describe the number of particles correctly. Let us plot $n_1(T,\mu) = N_1/V$ as a function of $T$ (Fig. 3.4.1).

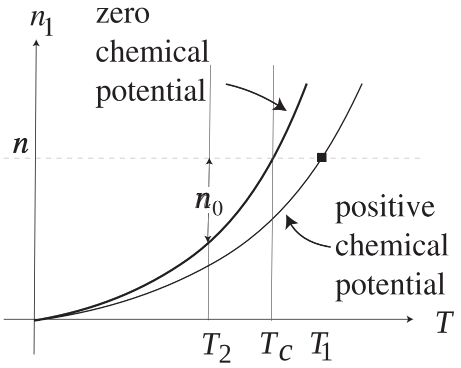


Fig. 3.4.1 Qualitative behavior of $n_1(T,\mu)$, the integral approximation = the number density of particles in excited states. If the temperature is higher than $T_c$ as $T_1$, we may choose a negative chemical potential such that $n = n_1$. If $T < T_c$ such as $T_2$, then $\mu = 0$, but then $n_0 > 0$ is possible and $n_1 + n_0 = n$ can be satisfied.

We must write the contribution of the ground state explicitly as

$$N = \frac{1}{e^{-\beta\mu}-1} + N_1(T,\mu). \tag{3.4.19}$$

Let us write the number densities with the lower case letters as $n_1 = N_1/V$, etc.
(i) If $n_1(T,0) > n = N/V$, $\mu(T) < 0$ is required, but then $N_0$ cannot be macroscopic, so we must choose a negative chemical potential such that $n_1 = n$ .
(ii) As we lower the temperature, to fulfill the equality $n_1 = n$, we must increase $\mu$. This is sketched in Fig. 3.4.1.
(iii) As we increase $\mu$, we hit $\mu = 0$ at $T = T_c$. Below this temperature $n_1 < n$, so (one particle) ground state is occupied by a macroscopic number of particles. This is called the *Bose-Einstein condensation*.[10][r]

Below $T_c$, the number density of the *condensate* is[q]

$$n_0(T) = n - n_1(T,0) = n\left(1 - \left(\frac{T}{T_c}\right)^{3/2}\right). \tag{3.4.20}$$

---

[10] This was discovered by Einstein; Bose has nothing to do with this discovery.

For small $T - T_c$ $n_0 \propto (T_c - T)$. To get $T_c$ we must solve $n_1(T_c, 0) = n$; from (3.4.18)

$$n_1(T_c,0) = \frac{1}{h^3}\int_0^\infty \frac{1}{e^{\beta p^2/2m}-1}4\pi p^2 dp = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} F_{3/2}(0) \tag{3.4.21}$$

with

$$F_{3/2}(0) = \frac{2}{\sqrt{\pi}}\int_0^\infty \frac{t^{1/2}}{e^t - 1}dt = \zeta(3/2) \simeq 2.61. \tag{3.4.22}$$

Therefore,[q]

$$T_c = \frac{h^2}{2\pi k_B m}\left(\frac{n}{2.61}\right)^{2/3} \propto \frac{n^{2/3}}{m}. \tag{3.4.23}$$

$1/m$ can be seen from the combination of $\beta p^2/2m$ in the Boltzmann factor.

Let us confirm[r] that in the pressure formula we may replace the summation (3.1.16) with the integration (3.1.22). The grand canonical partition function $\Xi_0$ for the lowest energy state $\varepsilon = 0$ is given by

$$\Xi_0 = \sum_{n=0}^{\infty} e^{n\beta\mu} = \frac{1}{1 - e^{\beta\mu}} = e^{-\beta\mu}\langle N_0\rangle. \tag{3.4.24}$$

Below $T_c$ $\mu$ is almost zero (really zero in the thermodynamic limit):[q]

$$\langle N_0\rangle = \frac{1}{e^{-\mu\beta} - 1} \simeq \frac{k_B T}{|\mu|}. \tag{3.4.25}$$

Because $\Xi_0 = 1/(1 - e^{\beta\mu}) = N_0 e^{1/N_0}$, the contribution of the ground state to pressure reads $P_0 \simeq (k_B T/V)\log N_0$. That is, it is $O[(\log N)/N]$. Therefore, we may use integration in macroscopic systems to obtain pressure (and internal energy).

Below $T_c$ $\mu = 0$ so thermodynamic quantities depends on $T$ in a clean fashion. This fact is worth memorizing. For example,[q]

$$E = \int d\varepsilon g D_t(\varepsilon)\frac{\varepsilon}{e^{\beta\varepsilon} - 1} = \frac{3}{2}k_B T V \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \zeta(5/2) \propto V T^{5/2}, \tag{3.4.26}$$

so

$$C_V \propto T^{3/2}. \tag{3.4.27}$$

For $T > T_c$ we must pay attention to $\mu(T)$ as well, so it is not very easy to demonstrate that $C_V$ is monotone decreasing (explicit calculation is needed; see the problem for this Chapter).

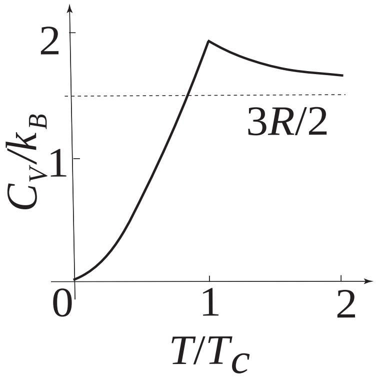


Fig. 3.4.2**q** Specific heat has a cusp at $T_c$. To see the behavior above $T_c$, we need a detailed calculation (see a problem at the end of the chapter).

The integral in (3.4.18) for $d \le 2$ dimensional space diverges, so $n_1 = n$ can always be satisfied by choosing an appropriate $\mu(T)$. Therefore, there is no Bose-Einstein condensation in low dimensional spaces.

When superfluidity was found (around 1938 explicitly by Kapitza[11]),[12] London suggested this is related to the Bose-Einstein condensation.

**r**To consider this problem, we must reflect on what 'superfluidity' means. Charged superfluidity is superconductivity. We know that supercurrent and the Meissner effect are two major characteristic features of superconductivity. We have already seen that the Bohr-van Leeuwen theorem should imply that the Meissner effect is a purely quantum effect. Furthermore, the rotation-vector potential analogy implies that there is a counterpart of the Meissner effect, the nonclassical rotational inertia (NCRI) for non-charged superfluid (Section 1.7). Therefore, as Leggett stresses, NCRI and supercurrent should be the characteristic features of superfluidity.

To exhibit NCRI the system must have a certain quantum coherence over a macro-

---

[11] Kamerlingh-Onnes found a singularity in the density in 1911, but Kapitza was the first to realize superfluidity: Nature, **141**, 74 (1938).

[12] [In 1911 Rutherford proposed his atom model. G. Mahler died; Amundsen reached the South Pole. In 1938 Hahn et al. discovered nuclear fission. Kemal Atatürk died; Carel Chapek died (who coined the word 'robot'). Anschluss; Kristallnacht.]

scopic range, because non-commutativity of the canonical variables is crucial. However, NCRI does not require the Bose-Einstein condensation.[13] For the existence of supercurrent, the state with flow must be metastable thermodynamically, so we must consider the barrier between different flow states. Again, some sort of quantum coherence is needed,[14] but its direct relation to the Bose-Einstein condensation does not exist.[15]

## 3.5 Phonons and photons

Photons and phonons are obtained through quantization of the systems that can be described as a collection of harmonic oscillators.[16] Possible energy levels for the $i$-th mode whose angular frequency is $\omega_i$[17] are $(n+1/2)\hbar\omega_i$. The canonical partition function of the whole systems is given by

$$Z(\beta)=\prod_i\left(\sum_{n_i}e^{-\beta(n_i+1/2)\hbar\omega_i}\right), \tag{3.5.1}$$

since no modes interact each other. Here, the product is over all the modes. The sum in the parentheses gives the canonical partition function for a single harmonic oscillator we have already encountered in (2.2.12). The canonical partition function may be rewritten as:

$$Z(\beta)=\prod_i\left(e^{-\beta\hbar\omega_i/2}\right)\Xi(\beta,0). \tag{3.5.2}$$

---

[13] Easy counterexamples exist. See A. Leggett, 'Topics in superfluidity and superconductivity' in Lecture Notes Phys **394** (1991).

[14] ⟪**1D superfluidity**⟫ See Eggel et al., "Dynamical Theory of Superfluidity in One Dimension," Phys. Rev. Lett.,**107**, 275302 (2011). The paper asserts superflow is essentially a dynamical phenomenon related to the suppression of quantum phase slip at low temperatures.

[15] A natural conclusion is that superfluidity and the Bose-Einstein condensation are caused in $^4$He liquid by a deeper common reason, the special topology of the many-body wave function.

[16] That is, the system whose Hamiltonian is quadratic in canonical coordinates (quantum mechanically in the corresponding operators).

[17] A system with a quadratic Hamiltonian may be described in terms of canonical coordinates (or corresponding operators) that makes the Hamiltonian diagonal. In other words, the system may be described as a collection of independent harmonic oscillators. The motion corresponding to each such harmonic oscillator is called a *mode*. If more than one modes have identical angular frequencies, modes cannot be uniquely chosen, but this does not cause any problem to us because partition functions need the system energies and their degeneracies only.

Here, we have used the formula

$$\Xi(\beta, 0) = \prod_i \left( \sum_{n=0}^{\infty} e^{-\beta n \hbar \omega_i} \right), \tag{3.5.3}$$

which may be obtained from (3.1.8) by setting $\varepsilon_i = \hbar\omega_i$, and $\mu = 0$. As long as we consider a single system, the total zero-point energy of the system $\sum_i \hbar\omega_i/2$ may be ignored by shifting the energy origin.[18] Therefore, the canonical partition function of the system consisting of photons or phonons may be written as $\Xi_{BE}(\beta, 0)$. That is, it is written as the grand partition function with a zero chemical potential.

The thermodynamic potential for the system consisting of photons or phonons is the Helmholtz free energy $A$ whose independent variables are $T$ and $V$, because the expected number $\langle n_i \rangle$ of phonons (photons) of mode $i$ is determined, if the temperature $T$ and the volume $V$ are given. Since $dA = -SdT - PdV$, we have $A = -PV$. That is, our observation $\log Z(\beta) = \log \Xi(\beta, 0)$ holds as a thermodynamic relation for a system that can be described by a collection of harmonic oscillators (as long as we ignore the zero-point energy). Thus, we may conclude that systems consisting of phonons or photons can be described consistently by the grand partition function with a zero chemical potential. For example, the pressure of the photon or phonon system can be computed immediately as we see below. However, do not understand this relation to indicate that the chemical potentials of photons and phonons are indeed zero; actually they cannot be defined. The relation is only a mathematical formal relation that can be sometimes useful.

Photons are super-relativistic particles in 3-space, so the relation between the one particle energy and the momentum (dispersion relation) can be written as $\varepsilon = c|\boldsymbol{p}|$, where $c$ is the speed of light. Therefore, the one particle translational energy density reads $D_t(\varepsilon) \propto \varepsilon^2$. Consequently, with the aid of the logic around (3.1.23), we have

$$PV = \frac{1}{3}E. \tag{3.5.4}$$

This was known to Boltzmann (of course before photons) from classical electrodynamics. This and thermodynamics tell us that $E \propto VT^4$ (the *Stefan-Boltzmann*

---

[18] However, if the system is deformed or chemical reactions occur, the system zero-point energy can change, so we must go back to the original formula with the total zero-point energy and take into account its contribution. We have already seen this when we computed the chemical equilibrium constant (just above (2.9.28)). Furthermore, for electromagnetic field, the change of the total zero-point energy may be observed as force. This is the *Casimir effect*.

*law*):[19] $dE = TdS - PdV$ implies

$$\left.\frac{\partial E}{\partial V}\right|_T = T\left.\frac{\partial P}{\partial T}\right|_V - P. \tag{3.5.5}$$

Let $E = Ve(T)$. Then

$$e(T) = \frac{1}{3}Te'(T) - \frac{1}{3}e(T) \tag{3.5.6}$$

or

$$Te'(T) = 4e(T). \tag{3.5.7}$$

That is, $e(T) \propto T^4$. Since $PV = -A \propto VT^4$, the entropy of photon gas must behave as $S \propto VT^3$.

Statistical mechanics allows us to compute the proportionality coefficient in $P \propto T^4$. The translational energy density $D_t(\varepsilon)$ of photons may be obtained as usual from

$$\int_0^\varepsilon D_t(\varepsilon')d\varepsilon' = \frac{4\pi}{3h^3}Vp(\varepsilon)^3, \tag{3.5.8}$$

where $p(\varepsilon)$ is the magnitude of momentum corresponding to energy $\varepsilon$, and is given by $p = \varepsilon/c$. Introducing this, we obtain

$$D_t(\varepsilon)d\varepsilon = 4\pi V\frac{\varepsilon^2}{c^3}d\varepsilon = 4\pi V\frac{\nu^2}{c^2}d\nu, \tag{3.5.9}$$

where $\nu$ is the frequency. The multiplicity $g$ of a photon is 2 (the polarization is perpendicular to the momentum)

$$gD_t(\varepsilon)d\varepsilon = 2 \times 4\pi V\frac{\nu^2}{c^3}d\nu. \tag{3.5.10}$$

The pressure may be obtained as

$$\begin{aligned} PV &= -k_BT\int_0^\infty \log(1-e^{-\beta h\nu})8\pi V\frac{\nu^2}{c^3}d\nu, & (3.5.11)\\ &= -\frac{(k_BT)^4V}{\pi^2(\hbar c)^3}\int_0^\infty \log(1-e^{-x})x^2dx = \frac{(k_BT)^4V}{3\pi^2(\hbar c)^3}\int_0^\infty \frac{x^3}{e^x-1}dx. & (3.5.12)\end{aligned}$$

---

[19] About Stefan and Boltzmann see D. Lindley, *Boltzmann's Atom, The great debate that launched a revolution in physics* (The Free Press, New York, 2001).

The integral can be performed ($= \pi^4/15$) to give

$$PV = \frac{4\sigma}{3c} VT^4, \tag{3.5.13}$$

where $\sigma$ is the *Stefan-Boltzmann constant*:

$$\sigma = \frac{\pi^2 k_B^4}{60\hbar^3 c^2} = 5.672 \times 10^{-8} \text{ kg/s}^3\text{K}^4. \tag{3.5.14}$$

We also get (from $dA = -SdT$)$_{\mathbf{q}}$

$$S = \frac{16\sigma}{3c} VT^3. \tag{3.5.15}$$

Therefore, the internal energy is given by the *Stefan-Boltzmann law*:

$$E = \frac{4\sigma}{c} VT^4 = 3PV. \tag{3.5.16}$$

From (3.5.12) we obtain

$$e(T) = \frac{(k_B T)^4}{\pi^2 (\hbar c)^3} \int_0^\infty \frac{x^3}{e^x - 1} dx. \tag{3.5.17}$$

Therefore, the energy per volume contained in the wavelength range $\lambda$ and $\lambda + d\lambda$ reads

$$e(\lambda, T) = \frac{8\pi hc}{\lambda^5} \frac{h\nu}{e^{\beta h\nu} - 1}. \tag{3.5.18}$$

This is *Planck's radiation law*.

**How Planck reached his law**

In 1900[20] Planck knew Wien's semiempirical formula (in terms of the Boltzmann distribution): asymptotically in the short wavelength limit

$$u(\lambda, T) = \frac{8\pi hc}{\lambda^5} e^{-\beta hc/\lambda}, \tag{3.5.19}$$

and an experimental result that $u(\lambda, T) \propto T$ in the long wavelength limit, which he knew only a few days before the German Physical Society meeting in October. It is very likely that Planck did not know the Rayleigh-Jeans formula

$$u(\lambda, T) = \frac{8\pi}{\lambda^4} k_B T \tag{3.5.20}$$

---

[20] [1900N: Britain annexed Republic of South Africa; Boxer Rebellion]

> that can be obtained by applying the equipartition of energy for harmonic oscillators.[21] During the meeting on October 19th after H. Rubens' presentation of his experimental results, Planck proposed (3.5.18) as an interpolation formula between the above two known results. Next morning, Rubens came to tell Planck that the formula fit well with experimental results over all the wavelengths. Thus, Planck's efforts to derive his formula began.[22]

We know (acoustic) phonons have the dispersion relation just as super-relativistic particles at the low energy end of the spectrum.[23] Therefore, the density of states for low frequency acoustic phonons must have the same form as (3.5.10):

$$gD_t(\nu) = 3 \times 4\pi V \frac{\nu^2}{c^3} d\nu, \tag{3.5.21}$$

where $c$ is the acoustic speed. '3' appears, because acoustic phonons have two transversal and one longitudinal waves.[24] Therefore, at low temperatures only long wavelength acoustic phonons are excited, detailed material structure should not matter and (3.5.21) is a good description. However, in contrast to photons, the number of the degree of freedom for phonons is finite for a finite volume (for acoustic phonons there are only $3N$, where $N$ is the number of unit cells in the crystal). Thus, Debye introduced the so-called *Debye approximation* for the acoustic phonon spectrum:

$$D(\nu) = \frac{9N}{\nu_D^3} \nu^2 \theta(\nu_D - \nu). \tag{3.5.22}$$

Here, $\nu_D$ is the *Debye frequency* and is determined by

$$\int_0^{\nu_D} D(\nu) d\nu = 3N. \tag{3.5.23}$$

---

[21] Rayleigh derived this formula without the exact proportionality constant in June, 1900. The proportionality constants were given, independently, by Einstein and Jeans in 1905.

[22] based on H. Ezawa's exposition "Quantum mechanics: from its birth to reconsideration of tis basis," in *Physics of the 20th century* (edited by H. Ezawa, Saiensu-Sha, 1998) [in Japanese].

[23] Notice this corresponds to the one-particle energy state density distribution for the ordinary ideal particle system.

[24] Here, for simplicity, we assume that the material is isotropic, but generally speaking the sound speed for transversal and longitudinal waves are different, so it is often used an effective sound speed $c$ defined by

$$3/c^3 = 1/c_L^3 + 2/c_T^3,$$

where $c_L$ is the speed of the longitudinal sound wave and $c_T$ that of the transversal sound wave.

The rest is quite parallel to the black-body radiation (if we ignore the existence of the frequency cutoff $\nu_D$). The internal energy reads

$$E(T) = \int_0^{\nu_D} \frac{h\nu D(\nu)}{e^{\beta h\nu} - 1} d\nu. \tag{3.5.24}$$

Therefore, the specific heat is given by

$$C_V = 3Nk_B \frac{3}{\nu_D^3} \int_0^{\nu_D} \frac{e^{\beta h\nu}}{(e^{\beta h\nu} - 1)^2} \left(\frac{h\nu}{k_BT}\right)^2 D(\nu) d\nu = 3Nk_B f(\Theta_D/T). \tag{3.5.25}$$

Here, $\Theta_D = h\nu_D/k_B$ is called the *Debye temperature*,[25] and $f$ is defined by

$$f(x) = \frac{3}{x^3} \int_0^x \frac{t^4 e^t}{(e^t - 1)^2} dt. \tag{3.5.26}$$

A major difference of (3.5.24) from the Stefan-Boltzmann law is the existence of the high-frequency cutoff $\nu_D$ due to the existence of the crystal lattice. However, if the temperature is sufficiently low ($T \ll \Theta_D$), no high frequency phonons can be significantly excited, so we may take $\nu_D \to \infty$ limit. Thus, at low temperatures $E \propto T^4$ and $C_v \propto T^3$ (*Debye's $T^3$ law*):

$$C_V \simeq 3Nk_B \frac{12\pi^4}{5} \left(\frac{T}{\Theta_D}\right)^3. \tag{3.5.27}$$

In the high temperature limit, the equipartition of energy implies

$$C_V \simeq 3Nk_B \frac{3}{\nu_D^3} \int_0^{\nu_D} D(\nu) d\nu = 3Nk_B, \tag{3.5.28}$$

which is the *Dulong-Petit law*. The discrepancy between the true specific heat $C_V(T)$ and the Dulong-Petit law gives the zero point energy $E_0$ of the solid under the harmonic approximation:[26]

$$\int_0^\infty [C_V(\infty) - C_V(T)] dT = E_0. \tag{3.5.29}$$

---

[25] **⟪Debye temperature is not a constant⟫** The Debye temperature (or frequency) is not a constant for a given material, but is a function of $T$ and $P$, because the frequency spectrum of a given material (even if it is pure) depends on them.

[26] This holds for any phonon spectrum, so no Debye approximation is required to show this.

The reader should have realized that in $D$-space:
(1) The existence of the Fermi level implies $C_V \propto T^{(D-1)/2}$ ($T^{D-1}$ superrelativistically) at low temperatures.
(2) Noninteracting bosons imply $C_V \propto T^{D/2}$ ($T^D$ superrelativistic case[27]).
(3) For any ideal gas $PV = (2/D)E$ ($(1/D)E$ superrelativistically[28]).
These are quite *universal*; they depend almost only on the general symmetry of particles and the spatial dimensionality.

Michael Fisher classifies statistical physics problems into two classes, fetish and fundamental.[29] To compute, e.g., the actual melting point is a typical fetish problem; it depends sensitively on materials details. This is an important problem for materials science that cashes on nonuniversality of materials (peculiarities of particular materials), but it is not a major question of statistical physics.

In this chapter we studied universal features of noninteracting systems. Thus, we studied fundamental questions of statistical physics. The universalities summarized above are, however, based on the universality of the building blocks of the systems. Thus, these universalities are straightforward manifestation of atomistic universality; therefore, we may say they are *trivial (or elementary) universalities.*[30] In Nature there are other kinds of universalities that emerge from truly many-body interactions such as critical phenomena and polymer solutions. These universalities are highly nontrivial and cannot easily be guessed or understood from the nature of the constituents of the system. In contrast to the trivial universalities, these gen-

---

[27] It is an easy exercise to see that the dispersion relation $\varepsilon \propto |\boldsymbol{p}|^s$ implies $C_V \propto T^{D/s}$.

[28] For the dispersion relation $\varepsilon \propto |\boldsymbol{p}|^s$ $PV = (s/D)E$.

[29] ⟪**The task of theory**⟫ M. Fisher, *Scaling, Universality and Renormalization Group Theory*, Lecture Notes in Physics **186** (1983) p4 says:

> What is the task of theory? It is worthwhile, when embarking on theory to have some viewpoint as to what theory is. There are different opinions on this subject. Some people feel the task of theory is to be able to calculate the results of any experiment one can do: they judge a theory successful if it agrees with experiment. That is not the way I look at a theory at all. Rather, I believe the task of theory is to try and understand the universal aspect of the natural world; first of all to identify the universal; then to clarify what they are about, and to unify and inter-relate them finally, to provide some insights into their origin and nature. Often a major step consists in finding a way of looking at things, a language for thinking about things — which need not necessarily be a calculational scheme.

[30] However, the reader should not automatically conclude that the case that looks to be described by a collection of free particles is always trivial. The quasiparticle in Landau's Fermi liquid theory is a counter example.

uine universalities are robust against modifying microscopic details except for some crucial parameters. The nontrivial universalities are important topics of Chapter 4.

## 3.6 Classical ideal gas — internal degrees of freedom

If the density is not very high or the temperature is not too low, gases behave as classical ideal gas. However, internal degrees of freedom of each gas molecule may not be classical.**r** We already know that in most cases the actual specific heat $C_V$ does not agree with the result due to the equipartition of energy. This is simply because excitation of many of the internal degrees of freedom requires energy quanta far greater than $k_BT$.

Let us first itemize internal degrees of freedom:
i) Each atom has a nucleus, and its ground state could have nonzero nuclear spin. This interacts with electronic angular momentum to produce *ultrafine structure*. The splitting due to this effect is very small, so for the temperature range relevant to gas phase we may assume all the levels are energetically equal. As we will see in the case of homonuclear diatomic molecules, nuclear spins could interfere with rotational degrees of freedom through quantum statistics, but otherwise we can simply assume that the partition function is multiplied by a constant $g$ = degeneracy of the nuclear ground state.
ii) Electronic degrees of freedom has a large excitation energy (of order of ionization potential ∼a few eV, so unless the ground state of the orbital electrons) is degenerate, we may ignore it. If the ground state is degenerate, then it could have a fine structure with an energy splitting of order a few hundred K[31], so we cannot simply assume that all the states are equally probable nor that only the ground slate is relevant.
iii) If a molecule contains more than one atom, it can exhibit rotational motion. The quantum of rotational energy is usually of order 10K.[32]

---

[31] For ground state oxygen ($^3P_2$) the splitting energy is about 200K.

[32] However, for $H_2$ it is 85.4K. For other molecules, the rotational quantum is rather small: $N_2$: 2.9K; HCl: 15.1K.

iv) Also such a molecule can vibrate. The vibrational quantum is of order 1000K.[33]

Notice that there is a wide temperature range, including the room temperature, where we can ignore vibrational excitations and can treat rotation classically (Fig. 3.6.1). Thus, equipartition of energy applied to translational and rotational degrees of freedom can explain the specific heat of many gases.

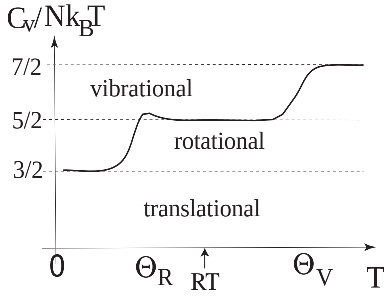


Fig. 3.6.1
The constant volume specific heat of diatomic gas. Note the local maximum due to the rotational contribution.**q**

Due to the indistinguishability of identical particles, if a diatomic molecule is *homonuclear* (that is $A$-$A$ type instead of $A$-$B$), rotation by 180 degrees gives the microstate indistinguishable from the original state. This does not happen if the nuclei are distinct (*heteronuclear* case). Therefore, we must discuss homo and heteronuclear diatomic molecules separately.[34]

First, let us look at the heteronuclear diatomic case. Diatomic molecules have rotational and vibrational degrees of freedom. The energy quantum of rotation $\varepsilon_r$ is $1 \sim 10$K (for $H_2$ it is exceptionally high: 85.4K). The energy quantum of vibration $\varepsilon_v$ is $\sim 10^3$ (for $H_2$ this is 6100K, again exceptionally high). Notice that $\varepsilon_r \ll \varepsilon_v$. This implies that the coupling between rotation-vibration must be very weak and can be ignored. Nonlinearity (*anharmonicity*) in vibration is also about the same order as the rotation-vibration coupling. Therefore, the Hamiltonian for the internal degrees of freedom reads

$$H = \frac{1}{2I}J^2 + \hbar\omega\left(\hat{n} + \frac{1}{2}\right), \tag{3.6.1}$$

where $I$ is the moment of inertia, $J$ the total angular momentum and $\hat{n}$ is the phonon number operator. Therefore, the partition function for the internal degrees of freedom reads

$$z_i = (2S_A + 1)(2S_B + 1)z_r z_v \tag{3.6.2}$$

---

[33] $N_2$ 3340K; $O_2$: 2260K; $H_2$: 6100K.

[34] Such complications should exist even for more complicated molecules, but quantum effects are significant only at low temperatures. Therefore, practically we need not worry such complications.

Here, $(2S_A+1)(2S_B+1)$ is due to the nuclear spins,

$$z_r = \sum_{J=0}^{\infty}(2J+1)e^{-\beta\hbar^2 J(J+1)/2I}, \tag{3.6.3}$$

and

$$z_v = \sum_{n=0}^{\infty} e^{-\beta\hbar(n+1/2)}. \tag{3.6.4}$$

$z_r$ may be written as

$$z_r = \sum_{J=0}^{\infty}(2J+1)e^{J(J+1)\Theta_r/T}, \tag{3.6.5}$$

where $\Theta_r = \hbar^2/2k_BI$. If $T \ll \Theta_r$, we may approximate

$$z_r \simeq 1 + 3e^{-2\Theta_r/T}. \tag{3.6.6}$$

The contribution of rotation to specific heat is

$$C_{rot} \simeq 3N\left(\frac{\hbar^2}{IT}\right)e^{-2\Theta_r/T}. \tag{3.6.7}$$

For $T \gg \Theta_r$, we may approximate the summation by integration (Large $J$s contribute, so we may approximate $J \simeq J+1$):

$$z_r \simeq 2\int_0^{\infty} dJ J e^{-J^2(\Theta_r/T)} = \frac{T}{\Theta_r}. \tag{3.6.8}$$

With the aid of the Euler-MacLaughlin summation formula[35] an accurate approximation is available:

$$z_r = \frac{T}{\Theta_r}\left\{1 + \frac{\Theta_r}{3T} + \frac{1}{15}\left(\frac{\Theta_r}{3T}\right)^2 + \frac{4}{315}\left(\frac{\Theta_r}{3T}\right)^3 + \cdots\right\}. \tag{3.6.9}$$

---

[35] ⟪**Euler-MacLaughlin summation formula**⟫

$$\sum_{m=0}^{\infty} f(m) = \int_0^{\infty} f(x)dx + \frac{1}{2}f(0) - \frac{1}{12}f'(0) + \frac{1}{720}f^{(3)}(0) + \cdots,$$

The formula is not necessarily a converging series. There is no guarantee that we can get a better approximation by retaining more terms. For example, try $\sum e^{-m^2}$.

The specific heat reads

$$C_{rot} \simeq k_B \left\{ 1 + \frac{1}{45}\left(\frac{\Theta_r}{3T}\right)^2 + \frac{16}{945}\left(\frac{\Theta_r}{3T}\right)^3 + \cdots \right\}. \tag{3.6.10}$$

The leading term is a consequence of the equipartition of energy. If $T$ is large, $\partial C_{rot}/\partial T < 0$, so the rotational specific heat has a peak and its maximum value is *larger* than the classical value $Nk_B$.

The vibrational partition function can be summed as

$$z_v = 1/2\sinh(\beta\hbar\omega/2). \tag{3.6.11}$$

For small $T$

$$z_v \simeq (1 + e^{-\beta\hbar\omega})e^{\beta\hbar\omega/2} \tag{3.6.12}$$

is enough. Consequently,

$$C_{vib} \sim k_B N \left(\frac{\hbar\omega}{k_B T}\right)^2 e^{-\beta\hbar\omega}. \tag{3.6.13}$$

Since $\Theta_r \ll \hbar\omega/k_B$, there is a wide range of temperature where only rotation contributes to the specific heat.

Next, let us study the homonuclear diatomic ideal gas.[r] The total wave function of the diatomic molecule reads

$$\psi = \psi_e \psi_r \psi_v \psi_n, \tag{3.6.14}$$

where $\psi_e$ is the electron wave function, $\psi_r$ the rotational wave function, $\psi_v$ the vibrational wave function, and $\psi_n$ the nuclear spin wave function. Two nuclei cannot be distinguished, so under their permutation $\mathcal{N}$ the total wave function $\psi$ must be symmetric (bosonic nuclei) or anti-symmetric (fermionic nuclei) (see the beginning of this chapter). We know

$$\begin{aligned} \mathcal{N}\psi_e &= \psi_e, & (3.6.15)\\ \mathcal{N}\psi_v &= \psi_v, & (3.6.16)\\ \mathcal{N}\psi_r &= (-1)^J\psi_r. & (3.6.17) \end{aligned}$$

The first two are simply because the electron or phonon clouds are not affected at all. The last line follows from the property of the spherical harmonics $Y_{Jm}$ (or recall

$s, p, d, f, \cdots$ electrons). Thus $\psi_n$ must also be an eigenstate of $\mathcal{N}$ (i.e., it must have a definite parity). The parity of $\psi_n$ depends on the nuclear spins: Let $\psi_A(s)$ be the spin state of nucleus $A$ with the $z$-component $s$. Then, eigenstates of $\mathcal{N}$ may be constructed as:
Symmetric case

$$\psi_n^{(+)} \propto \psi_A(s)\psi_B(s') + \psi_A(s')\psi_B(s). \tag{3.6.18}$$

Antisymmetric case

$$\psi_n^{(-)} \propto \psi_A(s)\psi_B(s') - \psi_A(s')\psi_B(s). \tag{3.6.19}$$

Suppose the nucleus has spin $S$. The total number of the distinguishable nuclear spin states is $(2S+1)^2$. If $s = s'$, only symmetric states are possible. There are $2S+1$ such 'diagonal' states. The remaining $2S(2S+1)$ states can have both symmetric and antisymmetric states. Therefore, there are $S(2S+1)$ distinct antisymmetric states,[36] and $(S+1)(2S+1)$ symmetric states.

Since the total symmetry of the wave function is constrained, we cannot consider rotational and nuclear wave functions independently.q Suppose nucleons are fermions, then antisymmetric (resp., symmetric) $\psi_n$ is possible only with even (resp., odd) $J$ states. Therefore, the rotational and nuclear partition function for fermions reads

$$z_{nr} = (2S+1)(S+1)z_r^{(o)} + S(2S+1)z_r^{(e)}, \tag{3.6.20}$$

where

$$z_r^{(o)} = \sum_{J=odd} (2J+1)e^{-J(J+1)\Theta_r/T}, \tag{3.6.21}$$

$$z_r^{(e)} = \sum_{J=even} (2J+1)e^{-J(J+1)\Theta_r/T}. \tag{3.6.22}$$

It is clear that for bosonic homonuclear diatomic molecules

$$z_{nr} = (2S+1)(S+1)z_r^{(e)} + S(2S+1)z_r^{(o)}. \tag{3.6.23}$$

In summary, the total partition function due to the internal degrees of freedom reads

$$z_i = z_e z_v z_{nr}. \tag{3.6.24}$$

At sufficiently high temperatures

$$z_r^{(e)} \simeq z_r^{(o)} \simeq \frac{1}{2} z_r. \tag{3.6.25}$$

---

[36] This is also ${}_{2S+1}C_2$.

Therefore, at sufficiently high temperatures

$$z_i = \frac{1}{2} z_e z_v z_r z_n. \tag{3.6.26}$$

The prefactor 1/2 is the *symmetry factor* to avoid double counting of geometrically indistinguishable states.

The above homonuclear diatomic molecules may be a good elementary quantum mechanics topic, which is no more very interesting (except in exams). However, the topic played an important role in determining whether protons are fermions or not.

**Is a proton a fermion?**
Although it was known that the proton spin was 1/2, before 1930 the spin-statistics relation was not known. For $H_2$ the rotational-nuclear partition function reads

$$z_{nr} = g_e z_r^{(e)} + g_o z_r^{(o)}. \tag{3.6.27}$$

If protons are fermions, $\beta = g_e/g_o = [(2S+1)S]/[(2S+1)(S+1)] = 1/3$, if bosons, its reciprocal, 3. In February 1927[37] Hund, analyzing the specific heat, found $\beta = 2$. Spectroscopically, Hori obtained $\beta = 1/3$. Dennison realized that there is no equilibrium between the singlet and triplet nuclear spin states in hydrogen gas (respectively called para hydrogen and ortho hydrogen) at low temperatures. Therefore, the nuclear-rotation partition function cannot be written as

$$Z_{nr} = z_{nr}^N = (g_e z_r^{(e)} + g_o z_r^{(o)})^N \tag{3.6.28}$$

but

$$Z_{nr} = (g_e z_r^{(e)})^{\rho N} (g_o z_r^{(o)})^{(1-\rho)N}, \tag{3.6.29}$$

where $\rho$ is the fraction of the even rotational angular momentum states. Dennison found $\rho = 1/4$ from the specific heat data. This implies that at sufficiently high temperatures where two nuclear spin states can change into each other $\beta = 1/3$ (on June 3, 1927). Then, later he realized that this implies that protons are fermions (on June 16, 1927).[38]

**Annealed and quenched systems**
The difference between (3.6.28) and (3.6.29) is very important in the study of a system under the influence of external fields or randomness in the system. Suppose the system depends on a parameter (or a field) $f$, and $f$ is a stochastic variable. For a fixed $f$, the partition function is given by $Z(f)$.

If $f$ varies sufficiently rapidly so that within the observation time $f$ samples its distribution almost evenly ('ergodically'), then the free energy we observe should be given by

$$A = -k_B T \log \langle Z(f) \rangle_f, \tag{3.6.30}$$

---

[37] [1927: This is the year Heisenberg proposed the uncertainty principle and matrix dynamics. Chiang Kai-shek set up a government in Nanjing; Mao moved to the Jinggang Mountains, Jiangxi]

[38] S. Tomonaga, *Spin the Spin* (World Scientific, 2004).

where $\langle\ \rangle_f$ is the average over $f$. However, if $f$ is frozen for each sample, but is different from sample to sample, then the free energy we get should be

$$A = -k_BT\langle \log Z(f)\rangle_f. \tag{3.6.31}$$

The former is the *annealed case* and the latter *quenched case*. The low temperature hydrogen gas was a quenched system.

Can we study the intermediate case with the aid of equilibrium statistical mechanics? No. Recall that an equilibrium state is a state where all the fast events have occurred, but no slow events have occurred yet.

Let us conclude this section with a summary of *polyatomic gases*. As in the case of diatomic gases, we can write

$$z_i = z_r z_v. \tag{3.6.32}$$

The nuclear contribution can be treated as a mere multiplicative factor as is discussed. Let the principal moments of inertia of the molecule be $I_1$, $I_2$ and $I_3$. Then the rotational Hamiltonian reads

$$H_{rot} = \sum_i \frac{L_i^2}{2I_i}, \tag{3.6.33}$$

where $L_i$ is the angular momentum around the $i$-th principal axis. We may treat the partition function classically,**q** so that

$$z_{rot} = \frac{(2k_BT)(\pi I_1I_2I_3)^{3/2}}{\hbar^3}. \tag{3.6.34}$$

Notice that $I_1I_2I_3$ is the determinant of the inertial tensor around the center of mass.

For a linear molecule $I_3 = 0$ and $I_1 = I_2 = I$, so that

$$z_{rot} = \frac{2Ik_BT}{\hbar^2}. \tag{3.6.35}$$

$z_v$ is the product of contributions from each normal mode. Often, vibrational degrees of freedom are all frozen, so we may ignore them.

## Exercises for Chapter 3

**3.1** [Fermions and bosons; the ultimate elementary problem]
There is a system with only three states with energies 0, $\varepsilon$ and $\varepsilon$ ($\varepsilon > 0$, i.e., excited states are degenerate). There are three identical particles without spin.
(1F) If the particles are fermions, write down the canonical partition function.
(2F) Find the probability of finding $N$ particles in the (one-particle) ground state.
(3F) Compute the average occupation number $N$ of the ground state.
(1-3B) Repeat the same problems assuming that the particles are bosons.
(4) Consider the high temperature limit. (UIUC Qual Spring00)

**3.2** [Elementary problem for boson systems]
There are 100 identical spinless bosons whose $s$-th one-particle state has an energy $E_s = s\varepsilon$ ($s \in \mathbf{N}$) and is described by a wave function $\phi_n(r)$ (normalized). These particles do not interact.
(1) How many microstates with the energy $4\varepsilon$ does the system have?
(2) When the system is in equilibrium with the particle reservoir (chemostat) of temperature $T$ and chemical potential $\mu$, on the average 99 particles occupy the one-particle ground state ($s = 0$), and one particle occupies the one-particle first excited state ($s = 1$). The other one-particle states are negligibly occupied, Find $\mu$ and $\beta = 1/k_BT$ in terms of $\varepsilon$.

**3.3** [Basic problem for quantum ideal gas: isothermal compression]
In a cylinder with a piston are $N$ identical particles. The temperature is maintained constant.
The fermion case:
(1) Suppose the system is maintained at $T = 0$, and the volume has been reversibly halved. What is the relation between the initial energy $e_i$ per particle and the final energy $e_f$ per particle?
(2) In the process described in (1) what is the ratio $P_f/P_i$, where $P_i$ is the initial pressure and $P_f$ the final pressure.
(3) Now, suppose the system is maintained at a positive temperature $T$. As in (1) we halve the volume of the system reversibly . How does the ratio $e_f/e_i$ change as a function of $T$? You may assume $T$ is sufficiently close to $T = 0$.
The boson case:
(4) Suppose the density of the condensate is positive at the initial temperature. After the volume is halved reversibly does the density of the condensate remain positive?
(5) Suppose $T = 0$ when the volume is reversibly halved. Find the ratio $P_f/P_i$ , where $P_i$ is the initial pressure and $P_f$ the final pressure.

**3.4** [Basic problem for quantum ideal gas: adiabatic free expansion]
In a thermally isolated cylinder with a piston is an ideal gas, whose initial tempera-

ture is $T_i$. The piston is pulled suddenly to increase the volume by 10%.
The fermion case: Suppose the ideal gas is fermionic.
(1F) Find the final pressure $P_f$ in terms of $P_i$, the initial pressure.
(2F) Which is correct, $T_i < T_f$, $T_i = T_f$ or $T_i > T_f$?
(3F) Suppose the initial temperature is $T = 0$. Express the final temperature $T_f$ approximately in terms of the Fermi temperature $T_F$ before the expansion.
The boson case: Suppose the ideal gas is bosonic.
(1B) Find the final pressure $P_f$ in terms of $P_i$, the initial pressure.
(2B) Suppose the initial temperature is sufficiently low and the condensate does not disappear by expansion. What is the final temperature $T_f$?
(3B) Suppose the initial temperature is less than $T_c$. After expansion, the final temperature becomes exactly $T_c$ (for the expanded system). Find the initial temperature $T_i$ in terms of the $T_c$ *before* expansion.

**3.5** [Basic problem for quantum ideal gas: adiabatic quasistatic expansion]
In a thermally isolated cylinder with a piston is an ideal gas, whose initial temperature is $T_i$ and initial pressure is $P_i$. The piston is pulled slowly to double the volume.
The fermion case: Suppose the ideal gas is fermionic.
(1F) Obtain the final pressure $P_f$ in terms of $P_i$.
(2F) What is the final temperature $T_f$, if $T_i = 0$?
(3F) More generally, obtain $T_f$ in terms of $T_i$.
The boson case: Suppose the ideal gas is fermionic.
(1B) Obtain the final pressure $P_f$ in terms of $P_i$.
(2B) Obtain $T_f$ in terms of $T_i$, assuming that the condensate does not disappear.
(4B) Let $N_{0i}$ be the initial number of particles in the condensate. Does the final number of particles $N_{0f}$ in the condensate increase or decrease?

**3.6** [Basic problem for quantum ideal gas: compression under constant internal energy]
In a cylinder with a piston is an ideal gas consisting of $N$ particles, whose initial temperature is $T_i$. The piston is pushed in slowly to halve the volume while removing thermal energy appropriately to keep the internal energy constant. Let $T_f$ be the final temperature.
I.The case of spinless bosons: assume that there is a Bose-Einstein condensate initially.
(1) Find the number of particles $N_0$ in the condensate before compression. You may use the critical temperature $T_c$.
(2) Which is true, $T_f < T_i$, $T_f = T_i$ or $T_f > T_i$?
(3) Does the number of particles in the condensate increase or decrease?
II. The case of spin 1/2 fermions.
(4) Find the final pressure $P_f$.

(5) Is there a minimum temperature ($> 0$) below which this process becomes impossible?
(6) Which is true, $T_f < T_i$, $T_f = T_i$ or $T_f > T_i$?

**3.7** [Qualitative properties of quantum ideal gases]
Assume the particles do not interact. Answer the following qualitative questions and give your justification for your answers. All the processes are quasistatic.
The boson case: there are $N$ bosons in a volume $V$.
(1B) The volume is increased under constant energy. Does the temperature decrease?
(2B) The volume is increased under constant entropy. Does the temperature decrease?
(3B) Can we decrease the volume while keeping the internal energy?
The fermion case: there are $N$ fermions in a volume $V$.
(1F) The volume is increased under constant energy. Does the temperature decrease?
(2F) The volume is increased under constant entropy. Does the temperature decrease?
(3F) Can we decrease the volume while keeping the internal energy?

**3.8** [Conversion of fermion into bosons]
There is an ideal fermi gas with the total energy 10 eV in an adiabatic container. The fermion particles are actually metastable and turn into bosons without adding any energy. Assume that the conversion is done quasistatically and adiabatically. Does the container explode? [UIUC qual]

**3.9** [Equation of state of ideal gases]
We know the relation between $PV$ and the internal energy does not depend on particle statistics.
(1) Is this still true for ideal gas mixtures?
(2) Compute $PV/E$ in d-space (this is already mentioned in the text).

**3.10** [Effective intereaction due to statistics]
Fig. 3.1.1 illustrates how we can intuitively understand the effective interactions between particles: compared with classical particles, between bosons there is an effective attraction, and between fermions there is an effective repulsion. Let us make this understanding slightly quantitative. Here, we proceed step by small step, reviewing elementary quantum mechanics.

We wish to consider a two-particle system in terms of canonical ensemble theory. The system Hamiltonian reads

$$H = \frac{\boldsymbol{p}_1^2}{2m} + \frac{\boldsymbol{p}_2^2}{2m}, \tag{3.P.1}$$

and the canonical partition function is

$$Z = Tr e^{-\beta H}, \tag{3.P.2}$$

where the trace is with respect to the microstates specified by two momenta $|\boldsymbol{p}, \boldsymbol{p}'\rangle$. To compute this trace semi-classically, we introduce a single-particle momentum state $|\boldsymbol{p}\rangle$.

(1) Express $|\boldsymbol{p}, \boldsymbol{p}'\rangle$ both for the boson and fermion cases in terms of single particle kets $|\boldsymbol{p}\rangle$. You may regard two momenta are distinct, but the obtained states must be properly normalized.

(2) Assuming that the system is in a sufficiently big box of volume $V$, find the position representation $\langle \boldsymbol{r}|\boldsymbol{p}\rangle$ (i.e., the wave function) of the momentum ket $|\boldsymbol{p}\rangle$.

(3) Let $\boldsymbol{r}_i$ be the position vector of the $i$-th particle. Find the position representation of $|\boldsymbol{p}, \boldsymbol{p}'\rangle$. [This is of course virtually the same question as (1).]

(4) For an $N$-particle system in the semi-classical limit, the calculation of trace in $Z$ may be performed as follows:

$$\begin{aligned} Tr &\to \frac{1}{N!}\int_{V^N} d\{\boldsymbol{r}_k\}\prod_{k=1}^{N}\langle \boldsymbol{r}_k|\cdots\prod_{k=1}^{N}|\boldsymbol{r}_k\rangle \qquad (3.P.3)\\ &= \frac{1}{N!}\int_{V^N} d\{\boldsymbol{r}_k\}\prod_{k=1}^{N}\langle \boldsymbol{r}_k|\left[\left(\sum_{\{\boldsymbol{p}_i\}}|\{\boldsymbol{p}_i\}\rangle\langle\{\boldsymbol{p}_i\}|\right)\cdots\left(\sum_{\{\boldsymbol{p}_i\}}|\{\boldsymbol{p}_i\}\rangle\langle\{\boldsymbol{p}_i\}|\right)\right]\prod_{k=1}^{N}|\boldsymbol{r}_k\rangle \qquad (3.P.4)\end{aligned}$$

If the volume is big enough, we should be able to replace the summation over momenta by integration over them. The replacement rule is

$$\sum_{\{\boldsymbol{p}_i\}} \to \frac{V^N}{h^{3N}}\int d\{\boldsymbol{p}_i\}. \tag{3.P.5}$$

Justify this for $N = 1$ in 1-space.

(5) Write $Z$ down using $h^{-3/2}e^{i\boldsymbol{r}\cdot\boldsymbol{p}/\hbar} = \langle \boldsymbol{r}_i|\boldsymbol{p}\rangle$. Beyond this point, let us simplify formulas by taking the $V \to \infty$ limit. You need not perform the integration.

(6) The outcome of (5) must have the following form:

$$\frac{1}{2h^6}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 d\boldsymbol{p} d\boldsymbol{p}' e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}[\cdots]. \tag{3.P.6}$$

Perform the integral in this expression and find $F$ in the following formula:

$$Z = \frac{1}{2h^6}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 d\boldsymbol{p} d\boldsymbol{p}'\, e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}F \tag{3.P.7}$$

(7) $F$ may be interpreted as the Boltzmann factor coming from the effective interaction originating from particle statistics. Sketch the potential ($\times\beta$) of this effective interaction for bosons and fermions.

**3.11** [Elementary low temperature formulas for fermions]
The following questions ask for standard elementary calculations, but you should do them once in your life.
(1)$_{\mathbf{r}}$ Obtain the chemical potential (the Fermi level) to order $T^2$ around $T = 0$D
(2) Obtain the pressure $P$ to order $T^2$ around $T = 0$.

**3.12** [Derivation of Maxwell's distribution]
Maxwell derived in his *Illustrations of the Dynamical Theory of Gases* (1860) the density distribution function $f(\boldsymbol{v})$ of the gas particle velocity.

Maxwell assumed that orthogonal components of the velocity are statistically independent. This implies that we may write

$$f(\boldsymbol{v}) = \phi_x(v_x)\phi_y(v_y)\phi_z(v_z), \tag{3.P.8}$$

where $\phi_x$, etc., are density distribution function for each component. Maxwell also assumed isotropy. Hence, $f$ is a function of $v^2 \equiv |\boldsymbol{v}|^2$, so we may regard $f(\boldsymbol{v}) \equiv F(v^2)$, and $\phi_x$'s do not depend on suffixes. Let us introduce $\psi(s^2) \equiv \phi_x(s)$. Then, the above functional equation reads

$$F(x + y + z) = \psi(x)\psi(y)\psi(z). \tag{3.P.9}$$

If $F$and $\psi$ are both once differentiable, we obtain

$$F'(x + y + z) = \psi(x)\psi(y)\psi'(z). \tag{3.P.10}$$

Setting $y = z = 0$, we have

$$F(x) = \psi(x)\psi(0)\psi(0), \quad F'(x) = \psi(x)\psi(0)\psi'(0), \tag{3.P.11}$$

so $F'(x)/F(x)$ must be a constant. This implies that the distribution is Gaussian.
(1) Is there any other solution? If we do not assume the differentiability of $F$ (that is, if we assume only the continuity of $F$), what do you get?[39]
(2) Since we know the result of equilibrium statistical mechanics, if the particle energy is $E$, then the distribution function is proportional to $e^{-\beta E}$. This is derived from the consistency of mechanics and thermodynamics. On the other hand, the above derivation of the Maxwell distribution uses only the statistical independence

---

[39] If we do not assume the continuity of $F$, there would be uncountably many solutions.

of the orthogonal components and its isotropy, and mechanics is never used.. Then, this seems to imply that Maxwell's logic determines the form of the kinetic energy $K$ in terms of velocity from statistically natural assumption + thermodynamics; at least $K \propto \boldsymbol{v}^2$ is concluded. This sounds incredible, even if thermodynamics is great. What is wrong? [Hint: think of relativistic case.]
[Comment] Maxwell himself did not like the above derivation we criticize here,[40] so he rederived the distribution a few years later. He this time used the detailed balance argument (as explained in the text). Pay due respect to Maxwell's sound instinct.

**3.13** [2-dimensional neutron system]
$10^{18}$ neutrons are confined in a square of edge length 1 m. If we regard this as a 2-dimensional system, estimate the needed temperature required for this system to be regarded a classical system.[41]

**3.14** [2-dimensional fermion system]
The density of translational states of a 2D fermion system confined in a volume (area) $V$ may be written as $c\ D(\varepsilon) = cV$ with a positive constant $c$.
(1) Find the chemical potential $\mu$ in terms of the number density $\rho$ and (inverse) temperature $\beta$.
(2) In the high density limit, we have $\mu \propto \rho$. Explain why this form is plausible.
(3) What is the classical limit? Does the obtained result consistent with the classical ideal gas result?

**3.15** [Quantum gas with internal degrees of freedom]
Let us consider a quantum gas consisting of $N$ particles. Individual particles have internal states consisting of two levels: the ground state and the non-degenerate excited state with energy $\varepsilon$ ($> 0$).
(1) Suppose the particles are fermions. How does the Fermi energy $\mu_F$ (i.e., the chemical potential) behave as a function of $\varepsilon$?
(2) Suppose the particles are bosons. How does the Bose-Einstein critical temperature $T_c$ depends on $\varepsilon$? Give a clear argument even if it is qualitative.

**3.16** [Zeemann splitting]
The outer shell of an ion has a magnetic moment $\mu_B$ of 1 Bohr magneton. In a magnetic field $B$ this outer shell state splits into two energy states with energies $E = E_0 \pm \mu_B B$. Let $n_u$ (resp., $n_d$) be the occupancy number of up-spin (resp., down-spin) states. Then the magnetization reads $M = \mu_B(n_u - n_d)$. You may ignore the electron-electron interactions.
(1) Find $\langle M\rangle$ and $\langle N\rangle$ ($N = n_u + n_d$) with the aid of the grand canonical formalism.

---

[40] However, even strict Pauli uncritically repeat the above argument in W. Pauli, *Thermodynamics and the Kinetic Theory of Gases* (edited by C. P. Enz), Section 25.

[41] cf ProblWS p176.

(2) Find the magnetization when the outer shell has one electron for each ion. Compare the result with the result of (1) for $\mu = E_0$.[42]D

**3.17** [Electron paramagnetism]
Due to the spin, each electron in a magnetic field $B$ (assumed to be pointing the $z$-direction) has the potential energy $\pm\tilde{\mu}B$. Let $D_t(\varepsilon)$be the one-particle translational density of states (however, the electrons may be in a crystal field, so we do not specify its form)D
(1) The magnetization $M$ of this system $M$ is the expectation of the magnetic moment due to electron spins. Express $M$ in terms of $D_t(\varepsilon \pm \tilde{\mu}B)$.
(2) Express the magnetic susceptibility $\chi$ in terms of $D_t'(\varepsilon)$, assuming that $\tilde{\mu}B$ is not too large.
(3) Obtain $\chi$ to order $T^2$ around $T = 0$ with the aid of $\log D_t(\varepsilon)$.

**3.18** [Do we have only to treat the ground state special below $T_c$?]
For a bose gas in 3-space we know the following integral expression is not always correct:

$$\langle N \rangle = \sum_{i=0}^{\infty} \langle \hat{n}_i \rangle = \int_0^\infty d\varepsilon\, D_t(\varepsilon). \tag{3.P.12}$$

It is because the expression ignores a large number of particles in the one-particle ground state. Thus, we are taught that if we count the number $N_0$ of the particles occupying the one-particle ground state and if we add this to $N_1$, then the number of particles in the system may be expressed correctly. However, ther may be the people who are not so convinced yet: why only ground state? Don't we have to consider the first excited state? Don't we actually have to perform the following calculation $\cdots$:

$$\frac{\langle N \rangle}{V} = \frac{1}{V}\langle \hat{n}_0 \rangle + \frac{1}{V}\langle \hat{n}_1 \rangle + \frac{1}{V}\int_0^\infty d\varepsilon\, D_t(\varepsilon). \tag{3.P.13}$$

Let us perform a slightly more honest calculation (to recognize clearly that Einstein is always correct!):
(1) Our energy coordinate convention is that the ground state is always 0: $\varepsilon_0 = 0$. Let us assume that the system is a cube of edge length $L$: $V = L^3$. The lowest excited one-particle state energy $\varepsilon_1$ as a function of $V$.
(2) Compare the occupation number of the one-particle ground state and the one particle first excited states (which is triply degenerate). That is, compute the ratio $\langle\langle \hat{n}_0 \rangle / (\langle \hat{n}_1 \rangle + \langle \hat{n}_2 \rangle + \langle \hat{n}_3 \rangle) = \langle \hat{n}_0 \rangle / 3\langle \hat{n}_1 \rangle$ for a very small negative chemical potential $\mu$[43] required by the Bose-Einstein condensation. How big is it as a function of $V$?

---

[42] UIUC QualFall 95

[43] which is not zero, because the system is finite.

(3) We just saw in (2) except for $\langle \hat{n}_0 \rangle$ other expectation values are not extensive. That is, the ground state is really special. Excited states cannot contribute an extensive quantity unless infinitely many of them are collected. Explain that the contribution of all the excited states may be obtained accurately by replacing the summation with integration (as usual).

**3.19** [Ideal boson gas slightly warmer than $T_c$]
Fig. 3.4.2 illustrates that the specific heat $C_v$ of the ideal Bose gas has a cusp. Let us demonstrate this. To compute $C_v$ we need the internal energy of the system. Let us compute it.
(1) What is the internal energy below $T_c$? (This is an easy question.)
(2) If we compute the internal energy assuming $\mu = 0$ and write its value as $E_0$, show

$$\frac{\partial E_0}{\partial \mu} \simeq \frac{3}{2} N_0(T), \tag{3.P.14}$$

where

$$N_0(T) \equiv N(T,0) = \frac{V}{h^3} \int_0^\infty \frac{1}{e^{\beta p^2/2m} - 1} 4\pi p^2 dp. \tag{3.P.15}$$

Therefore, for $T$ ($> T_c$) we could approximate the true internal energy at $T$ as $E(T) = E_0 + (3/2)N_0(T)\mu$. This implies that to obtain $E$ as a function of $T$, we need $\mu$ as a function of $T$. To this end let us write the number of particles for $T > T_c$ ($\mu < 0$) as

$$N = N_0(T) + \frac{V}{h^3} \int_0^\infty \left\{ \frac{1}{e^{\beta(p^2/2m - \mu)} - 1} - \frac{1}{e^{\beta p^2/2m} - 1} \right\} 4\pi p^2 dp. \tag{3.P.16}$$

(3) Show that we may approximate the second term of (6.3.123) as

$$\int_0^\infty \left\{ \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} - \frac{1}{e^{\beta\varepsilon} - 1} \right\} \sqrt{\varepsilon} d\varepsilon \simeq k_B T \mu \int_0^\infty d\varepsilon \frac{1}{\sqrt{\varepsilon}(\varepsilon + |\mu|)} = -\pi k_B T \sqrt{|\mu|}. \tag{3.P.17}$$

Do not forget that $\mu < 0$. [This allows us to obtain $\mu$ in terms of $N_0(T)$ which is obtainable from (6.3.122) as a function of $T$.]

**3.20** [Bose-Einstein condensation in a harmonic trap]
Let us consider an ideal bose gas consisting of $N$ particles confined in a 3D harmonic potential.[44] It is hard to treat this in terms of the canonical ensemble, so we discuss

---

[44] Actually, BEC is observed in a collection of Rb atoms confined in a (not spherically symmetric) 3D harmonic potential.

this with the aid of the grand canonical theory; if $N$ is larger than $10^3$, then $\log N/N$[45] is not large, so this approach must not be bad.
(1) Suppose the angular frequency of the trapped boson is $\omega_t$. Find the density $D(\varepsilon)$ of one-particle state as a function of energy $\varepsilon$. Measure the energy from the ground state and ignore the zero-point energy.
(2) Find the number of particles $N_1$ in the non-condensate as a function of the chemical potential. Show that the integral (or $N_1$) is bounded from above in 3-space (no explicit integration required). Thus, Bose-Einstein condensation is possible in this system.
(3) The number of particles occupying the one-particle ground state approaches zero as

$$N_0(T) = N\left(1 - \left(\frac{T}{T_c}\right)^{\gamma}\right), \tag{3.P.18}$$

when $T \nearrow T_c$. Find $\gamma$.
(4) Find $T_c$ as a function of $N$. For $N = 3000$, and $\omega_t = 10^3$ rad/s estimate $T_c$. (Use $\zeta(3) = 1.2020569031595\cdots$.)
(5) If a 2D harmonic potential can trap 2D Bose gas,[46] can we observe Bose-Einstein condensation?

**3.21** [Expanding universe]
At present, the cosmic background radiation is at 3 K. Suppose the volume of the universe doubles adiabatically. What is the temperature of the cosmic background radiation after this expansion?

**3.22** [Specific heat of hydrogens]
Consider a 1 mole of ideal gas at 10 K consisting of pure HD, pure HT or pure DT. Whose specific heat $C_V$ is the largest? Give your answer without detailed computation. You may assume that the length of the chemical bonds are all the same.

**3.23** [Internal degree of freedom of heavy hydrogen]
The potential energy function describing the chemical bond in a heavy hydrogen $D_2$ may be approximately described as

$$\phi(r) = \varepsilon\left[e^{-2(r-d)/a} - 2e^{-(r-d)/a}\right], \tag{3.P.19}$$

where $\varepsilon = 7 \times 10^{-19}$ J, $d = 8 \times 10^{-11}$ m and $a = 5 \times 10^{-11}$ m.
(1) Evaluate the smallest energy required to excite the rotational motion, and estimate the temperature $T_r$ for which the rotation starts to contribute significantly.

---

[45] It is emphasized again that the error is not of order $\sqrt{N}/N = 1/\sqrt{N}$, but $\log N/N$.

[46] This is virtually realized on graphene.

(2) Evaluate the smallest energy required to excite the vibrational motion, and estimate the temperature $T_rv$ for which the vibration starts to contribute significantly.

**3.24** [Computation of inertial moment tensor]
Obtain the moment of inertia tensor for $CH_3D$ around its center of mass, and compute its rotational partition function classically. [You have only to state your strategy without actually estimating the components of the tensor.]

# Chapter 4

# Interacting Fluid Systems

If there are interactions among the microscopic constituents of the system, structurally different phases with bewildering diversity become possible. Materials constants such as specific heat, elastic moduli, various susceptibilities sensitively depend on details of interactions. However, here, we are not so interested in materials diversity; the reader must clearly recognize the distinction between statistical physics and materials science. We concern about universal features and tools and concepts that are generally useful to understand equilibrium systems made of numerous interacting constituents. Thus, imperfect gases, liquids, magnets, etc., are discussed only to illustrate useful general ideas and tools.[1] We maximally avoid fetishism. In this chapter some tools to discuss classical many body systems are explained.

---

[1] For example, many-body problems in nuclear physics utilizes tools inspired by methods to study liquids.

## 4.1 Virial expansion of imperfect gas

A collection of identical classical particles interacting via a certain potential $U$ may be described by the following Hamiltonian

$$H = \sum_i \frac{p_i^2}{2m} + U(\{\boldsymbol{r}_i\}), \tag{4.1.1}$$

where $m$ is the particle mass and $\boldsymbol{r}_i$ the position vector of the $i$-th particle. It is often assumed that $U$ may be decomposed into binary interactions as

$$U = \sum_{\langle i,j \rangle} \phi(\boldsymbol{r}_{ij}), \tag{4.1.2}$$

where $\langle i,j \rangle$ denotes the pair (disregarding the order) of particles $i$ and $j$, $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$, and $\phi$ is the potential of the binary interaction. If we discuss a spherically symmetric interaction, then $\phi_{ij} \equiv \phi(\boldsymbol{r}_{ij}) = \phi(|\boldsymbol{r}_{ij}|)$ with a slight abuse of the function symbol.

How realistic is the binary interaction description? If the molecule is not simple and if the phase is dense, it is known that three-body interactions[2] are very important. However, it is also known that we could devise an effective binary interaction[3] that incorporates approximately the many-body effects. In this case the binary potential parameters are fitting parameters to reproduce macroscopic observables (thermodynamic and scattering data).[4]

---

[2] Here, 'three-body interaction' does not imply simultaneous binary interactions among three particle, but a genuine three-body interaction in the sense that even the intereaction between two particles is modified by the presence of the third particle.

[3] Later we will discuss the potential of mean force. Do not confuse this effective potential and the effective binary interaction being discussed here. In the present case, the 'effectiveness' implies to write truly many-body interactions (approximately) in terms of binary interactions. The potential of mean force we will discuss later is the potential of the effective force between a particular pair of particles surrounded by many other interacting particles.

[4] **⟪How good is the effective two-body interaction potential?⟫** The two-body correlation may be fitted with an effective binary interaction, but generally, we cannot fit the three-body correlation function. See, for example, an experimental work: C. Russ, M. Brunner, C. Bechinger, and H. H. von Grünberg, "Three-body forces at work: Three-body potentials derived from triplet correlations in colloidal suspensions," Europhys. Lett. **69**, 468 (2005). How good is the *ab initio* quantum mechanically obtained binary potential? For noble gases if we take into account the triple-dipole interaction, fairly good results for gas-liquid coexistence curve seems obtained. See A. E. Nasrabad and U. K. Deiters, "Prediction of thermodynamic properties of krypton by Monte Carlo simulation using *ab initio* interaction potentials," J. Chem. Phys. **119**, 947 (2003).

We already know from (2.2.7) that the classical canonical partition function $Z$ can be written as a product of the classical ideal gas partition function $Z_{ideal}$ and the configurational partition function $Q$, where

$$Q = \left\langle e^{-\beta \sum_{\langle i,j\rangle} \phi_{ij}} \right\rangle_V , \tag{4.1.3}$$

$V$ is the system volume, and $\langle * \rangle_V$ is the average (2.2.10) over the configuration space. If statistical mechanics is correct, all the phases (for which quantum effects may be ignored), gas phase, liquid phase, solid phase, etc., must be in $Q$.

**Isotope effects**

Isotope effects show up in spin-statistics relation and in mass difference. The spin-statistics relation affects only low temperature properties of diatomic molecules. Therefore, the effect due to mass difference is the only remaining isotope effect in most cases. There are two effects, modifying the de Broglie wave length and modifying the interaction potential (through changing the effective mass of electrons).

The ionic potential is of the order of 10 eV, so H-D mass difference could affect it and modify the interaction potential to the extent that cannot be totally ignored relative to $k_BT$. In any case, however, unless the H-D effect is involved, isotope effects are usually very small for equilibrium properties around room temperature. For low molecular weight compounds, the isotope effect is at most 1 or 2 K for phase transition temperatures.[5] However, the replacement of H with D in polymers can have a large effect. For example, the phase separation temperatures could change by 10 K. Polystyrene and D-polystyrene melts cannot mix. Thus, isotope effect may be amplified by many-body effects. Incidentally, heavy water disrupts the spindle to inhibit cell division, so seeds cannot germinate with heavy water; also it can cause male infertility.

The difference in de Broglie thermal wavelength can change the structure of condensed phases. For example, even if the interaction potential is the same, lighter isotopes tend to be delocalized.[6] In classical statistical mechanics, configurational and kinetic parts are separated, so unless there is a mass effect on the interaction potential, no isotope effect can be explained. Therefore, without modifying the potential to take account of the delocalization effect, classical statistical mechanics cannot explain the effect of de Broglie wavelength difference.

In terms of the configurational partition function $Q$, we can write (cf. $-V\partial A/\partial V = n\partial A/\partial n = PV$)

$$\frac{PV}{Nk_BT} = 1 - n\frac{\partial}{\partial n}W, \tag{4.1.4}$$

---

[5] For example, even for water the effect is small: the triple point: $H_2O$ is 0.01°C, $D_2O$ is 3.82°C; boiling point (at 1atm): $H_2O$ is 100°C, $D_2O$ is 101.42°C.

[6] A. Cunsolo, D. Colognesi, M. Sampoli, R. Senesi, and R. Verbeni, “Signatures of quantum behavior in the microscopic dynamics of liquid hydrogen and deuterium,” J. Chem. Phys. **123**, 114509, (2005). The paper observes that ordinary hydrogen delocalizes more than heavy hydrogen in the liquid phase.

where $W = (1/N)\log Q$ and $n = N/V$ (the number density). For gas phases, our main task is to obtain the *virial expansion* of the equation of state:[7]

$$\frac{PV}{Nk_BT} = 1 - \sum_{k=2}^{\infty}\frac{k}{k+1}\beta_k n^k, \tag{4.1.5}$$

$$= 1 + B(T)n + C(T)n^2 + D(T)n^3 + \cdots, \tag{4.1.6}$$

where $B, C, D, \cdots$ are called *virial coefficients.* The systematic calculation of virial coefficients is not very simple, but certain general useful ideas worth remembering are used: Mayer's $f$, cumulant expansion, 1-PI diagrams, etc.[8]

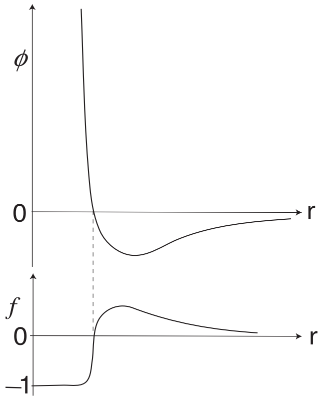


Fig. 4.1.1
Sketch of Mayer's $f$. It is short-ranged and usually between 1 and $-1$.

Expanding $W$ or $Q$ in terms of the number density is roughly equivalent to expanding it in terms of interactions. Unfortunately, the binary interaction potential has a hard core due to the Pauli exclusion principle, so the binary interaction potential $\phi$ is not bounded. Therefore, we need something smaller to facilitate our expansion. Mayer introduced *Mayer's f-function*:

$$f(r) = e^{-\beta\phi(r)} - 1. \tag{4.1.7}$$

In terms of $f$ we can write

$$Q = \left\langle \prod_{i<j} e_L^{f_{ij}} \right\rangle_V . \tag{4.1.8}$$

---

[7] This form was introduced by Kammerlingh-Onnes in 1901.

[8] Except for somewhat simplified notations, the exposition here is based on R. Abe, *Statistical Mechanics* (University of Tokyo Press) [in Japanese].

where $e_L^x = 1 + x$ is the *linearized exponential function.*[9]

**Elementary calculation of the second virial coefficient**[q]
Before jumping into a systematic calculation, let us compute the second virial coefficient in an elementary fashion.

$$\begin{aligned} Q &= \frac{1}{V^N}\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 \cdots d^3\boldsymbol{r}_N \prod_{i<j}(1+f_{ij}), && (4.1.9)\\ &= \frac{1}{V^N}\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 \cdots d^3\boldsymbol{r}_N \left(1+\sum_{i<j} f_{ij}+\right), && (4.1.10)\\ &= 1+\frac{1}{V^2}\binom{N}{2}\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 f_{12}+\cdots = 1+\frac{1}{V}\binom{N}{2}\int d^3\boldsymbol{\rho} f(\boldsymbol{\rho})+\cdots && (4.1.11) \end{aligned}$$

At the last step the integration variables have been switched from $\boldsymbol{r}_1, \boldsymbol{r}_2$ to $\boldsymbol{r}_1, \boldsymbol{\rho} = \boldsymbol{r}_2 - \boldsymbol{r}_1$. The integration over $\boldsymbol{r}_1$ gives $V$. Therefore,

$$W = \frac{1}{2}n\int d^3\boldsymbol{\rho} f(\boldsymbol{\rho}). \tag{4.1.12}$$

Comparing this with (4.1.6), we obtain

$$B(T) = -\frac{1}{2}\int d^3\boldsymbol{\rho} f(\boldsymbol{\rho}) = \frac{1}{2}\int d^3\boldsymbol{\rho}\,(1-e^{-\beta\phi(\boldsymbol{\rho})}). \tag{4.1.13}$$

If the binary potential is a hard core (of diameter $\sigma$) + short-ranged attractive tail,[r]

$$f(\rho) = \begin{cases} -1 \text{ for } \rho < \sigma, \\ -\beta\phi(\rho) \text{ for } \rho \geq \sigma, \end{cases} \tag{4.1.14}$$

so we get

$$B(T) = \frac{2}{3}\pi\sigma^3 - \frac{a}{k_BT}, \tag{4.1.15}$$

where we have introduced

$$a = -\frac{1}{2}\int_\sigma^\infty \phi(\rho)4\pi\rho^2 d\rho, \tag{4.1.16}$$

which is finite, since we have assumed a short-ranged attraction.

We need $\log Q$ instead of $Q$ itself. A general technique to study the logarithm of the generating function is the *cumulant expansion.* Let $X$ be a stochastic variable with all the moments well defiend. $\langle e^{\theta X}\rangle$ is the (moment) *generating function*:

$$\langle e^{\theta X}\rangle = 1 + \sum_{n=1}^\infty \frac{\theta^n}{n!}\langle X^n\rangle\,, \tag{4.1.17}$$

---

[9] The reason to introduce $e_L$ is to use cumulant expansion explained below as conveniently as possible later.

where $\theta$ is a real number. *Cumulants* are introduced as$_{\mathbf{r}}$

$$\log\langle e^{\theta X}\rangle = \sum_{n=1}^{\infty} \frac{\theta^n}{n!}\,\langle X^n\rangle_C = \left\langle e^{\theta X} - 1\right\rangle_C. \tag{4.1.18}$$

The second equality is for the mnemonics sake. $\langle X^n\rangle_C$ is called the $n$-th order cumulant. To compute cumulants, we must relate them to the ordinary moments. It is easy to extend the definition of cumulants to multivariate cases.

The following two properties are worth remembering:
(i) A necessary and sufficient condition for $X$ to be Gaussian is $\langle X^n\rangle_C = 0$ for all $n \geq 3$.
(ii) If $X$ and $Y$ are independent stochastic variables, then $\langle X^n Y^m\rangle_C = 0$ for any positive integers $n$ and $m$.
(i) is obvious from an explicit calculation. (ii) is obvious from $\langle e^{\alpha X+\beta Y}\rangle = \langle e^{\alpha X}\rangle\langle e^{\beta Y}\rangle$. (ii) is crucial in our present context.

We must have an explicit formulas for multivariate cases, but this extension becomes almost trivial with the aid of a clever notation (*Hadamard's notation*):[10] Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $D$-dimensional vectors. Then, we write

$$\boldsymbol{A}^{\boldsymbol{B}} = \prod_{i=1}^{D} A_i^{B_i}. \tag{4.1.19}$$

For (nonnegative) integer $D$-vector $\boldsymbol{N} = (N_1, \cdots, N_D)$

$$\boldsymbol{N}! = \prod_{i=1}^{D} N_i!. \tag{4.1.20}$$

With the aid of these notations, the *multivariate Taylor expansion* just looks like the one-variable case:

$$f(\boldsymbol{x} + \boldsymbol{x}_0) = \sum_{\boldsymbol{M}} \frac{1}{\boldsymbol{M}!} f^{(\boldsymbol{M})}(\boldsymbol{x}_0), \tag{4.1.21}$$

where $\boldsymbol{x}$ is a $D$-vector and

$$f^{(\boldsymbol{M})}(\boldsymbol{x}) = \left(\frac{\partial}{\partial \boldsymbol{x}}\right)^{\boldsymbol{M}} f(\boldsymbol{x}) \tag{4.1.22}$$

---

[10] The notation is standard in partial differential equation theory.

and the summation is over all the nonnegative integer $D$-vector $\boldsymbol{M}$: $\boldsymbol{M} \in \boldsymbol{N}^D$. The *multinomial theorem* reads

$$(x_1 + \cdots + x_D)^n = \sum_{\boldsymbol{K}} \frac{n!}{\boldsymbol{K}!} \boldsymbol{x}^{\boldsymbol{K}}, \tag{4.1.23}$$

where the summation is over all nonnegative component vector $\boldsymbol{K}$ such that $K_1 + \cdots + K_D = \boldsymbol{K} \cdot \mathbf{1} = n$. Here, $\mathbf{1} = (1, \cdots, 1)$ (there are $D$ 1's).

The *multivariate cumulants* are defined as (see (4.1.18))

$$\log \left\langle e^{\boldsymbol{\theta} \cdot \boldsymbol{X}} \right\rangle = \left\langle e^{\boldsymbol{\theta} \cdot \boldsymbol{X}} - 1 \right\rangle_C . \tag{4.1.24}$$

Actually, what we need later is only

$$\langle \boldsymbol{X}^{\boldsymbol{M}} \rangle_C = \langle \boldsymbol{X}^{\boldsymbol{M}} \rangle + \cdots , \tag{4.1.25}$$

where $\cdots$ has all the terms with extra insertion of $\rangle\langle$ (i.e., the terms decomposed into the product of moments).

**Cumulant in terms of moments**
Using

$$\log(1 + x) = \sum_n (-1)^{n-1} \frac{x^n}{n}, \tag{4.1.26}$$

we expand the LHS of (4.1.24) as

$$\log \left( 1 + \sum_{\boldsymbol{N} \neq 0} \frac{\boldsymbol{\theta}^{\boldsymbol{N}}}{\boldsymbol{N}!} \left\langle \boldsymbol{X}^{\boldsymbol{N}} \right\rangle \right) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \left\{ \sum_{\boldsymbol{N} \neq 0} \frac{\boldsymbol{\theta}^{\boldsymbol{N}}}{\boldsymbol{N}!} \left\langle \boldsymbol{X}^{\boldsymbol{N}} \right\rangle \right\}^n . \tag{4.1.27}$$

To expand the RHS we use (4.1.23). Therefore, the RHS of (4.1.27) becomes

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{\boldsymbol{K} \cdot \mathbf{1} = n} \frac{n!}{\boldsymbol{K}!} \left[ \frac{\boldsymbol{\theta}^{\boldsymbol{N}}}{\boldsymbol{N}!} \left\langle \boldsymbol{X}^{\boldsymbol{N}} \right\rangle \right]^{\boldsymbol{K}} . \tag{4.1.28}$$

Here, $[\cdots]^{\boldsymbol{K}}$ may be a slight abuse of the notation, but now each component of $\boldsymbol{K}$ denotes how many $\boldsymbol{\theta}^{\boldsymbol{N}}$ terms appear. This double summation over $n$ and $\boldsymbol{K}$ can be rewritten as

$$-\sum_{\boldsymbol{K}} (\boldsymbol{K} \cdot \mathbf{1} - 1)! (-1)^{\boldsymbol{K} \cdot \mathbf{1}} \frac{1}{\boldsymbol{K}!} \left[ \frac{\boldsymbol{\theta}^{\boldsymbol{N}}}{\boldsymbol{N}!} \left\langle \boldsymbol{X}^{\boldsymbol{N}} \right\rangle \right]^{\boldsymbol{K}} . \tag{4.1.29}$$

This should be compared with (4.1.24). Hence, we obtain

$$\left\langle \boldsymbol{X}^{\boldsymbol{M}} \right\rangle_C = -\boldsymbol{M}! \sum_{\sum_i K_i \boldsymbol{N}_i = \boldsymbol{M}} \left( \sum_i K_i - 1 \right)! (-1)^{\sum_i K_i} \frac{1}{\prod_i K_i!} \prod_i \left[ \frac{\left\langle \boldsymbol{X}^{\boldsymbol{N}_i} \right\rangle}{\boldsymbol{N}_i!} \right]^{K_i}. \tag{4.1.30}$$

Here, the summation is over all possible decomposition of $\boldsymbol{M}$ into nonzero nonnegative integer vectors $\boldsymbol{N}_i$ with multiplicity $K_i$.

We formally apply (4.1.24) to (4.1.8):

$$\log Q = \log \left\langle \prod_{i<j} e_L^{f_{ij}} \right\rangle_V = \sum_{\boldsymbol{M} \neq 0} \frac{1}{\boldsymbol{M}!} \left\langle \boldsymbol{f}^{\boldsymbol{M}} \right\rangle_C. \tag{4.1.31}$$

The cumulants can be expressed in terms of the moments $\langle \boldsymbol{f}^{\boldsymbol{N}} \rangle_V$, but thanks to the definition of $e_L$, no moment containing the same $f$ more than once shows up (i.e., the components of $N$ are 0 or 1).

Let us itemize several key ideas helpful to express cumulants in terms of moments. (1) There are many cumulants or moments, so a *diagrammatic expression* of $\boldsymbol{f}^{\boldsymbol{M}}$ is advantageous. In these diagrams, particles correspond to the vertices and $f_{ij}$ is denoted by a line connecting two vertices corresponding to particles $i$ and $j$ (See Fig. 4.1.2 for illustrations).

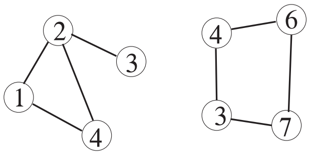


Fig. 4.1.2 The left diagram corresponds to $f_{12}f_{23}f_{14}f_{24}$, and the right one to $f_{34}f_{46}f_{67}f_{73}$.

There are many diagrams in these days, but the theory we are discussing here is the first systematic use of diagrammatics.

(2) If a diagram consists of two disconnected parts, then the particle positions can be changed freely independently, so the corresponding cumulants vanish. The same is true for diagrams that can be decomposed into disjoint pieces by removing one particle (one vertex). This is because the average of $\boldsymbol{f}^{\boldsymbol{M}}$ is the average over the relative coordinates between particles; if a diagram has a hinge around which two parts can be rotated independently, the interparticle vectors belonging to these two parts are statistically independent. Therefore, only the *1-PI diagrams* (= one-particle irreducible diagram; a diagram that does not decompose into disjoint parts by removing one vertex) matter.

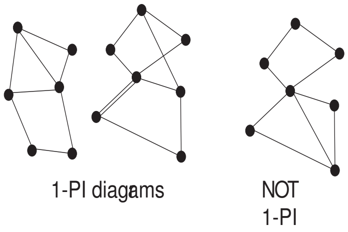


Fig. 4.1.3 1-PI diagrams and a non 1-PI diagram

(3) If $\boldsymbol{M}$ is a vector whose components are 0 or 1 and gives a 1-PI diagram, then we may identify $\langle \boldsymbol{f}^{\boldsymbol{M}} \rangle_C$ and $\langle \boldsymbol{f}^{\boldsymbol{M}} \rangle_V$ (as already announced) in the large volume limit. This may be understood as follows. Suppose the diagram corresponding to $\langle \boldsymbol{f}^{\boldsymbol{M}} \rangle$ has $m$ particles. Then, (let us call these $m$ particles $1, \cdots, m$)

$$\langle \boldsymbol{f}^{\boldsymbol{M}} \rangle_V = \frac{1}{V^m} \int d\boldsymbol{r}_1 \cdots d\boldsymbol{r}_m \boldsymbol{f}^{\boldsymbol{M}}. \tag{4.1.32}$$

The integral is of the order of $V$, because we may place the cluster at any position in the space. Therefore, the moment is of order $1/V^{m-1}$. Consider a decomposition $\boldsymbol{M}_1 + \boldsymbol{M}_2 = \boldsymbol{M}$.

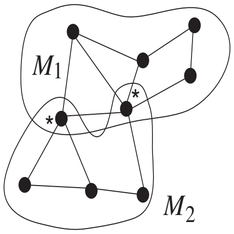


Fig. 4.1.4 A decomposition of a 1-PI diagram into two subdiagrams results in at least two common vertices (with $*$) shared by the subdiagrams. In this example, the edge between these shared vertices belong to $M_1$.

The original diagram is 1-PI, so the subdiagrams corresponding to $\boldsymbol{M}_1$ (containing $m_1$ vertices) and $\boldsymbol{M}_2$ (containing $m_2$ vertices) share at least two vertices. Therefore, $m_1 + m_2 - m \geq 2$ and

$$\langle \boldsymbol{f}^{\boldsymbol{M}_1} \rangle_V \langle \boldsymbol{f}^{\boldsymbol{M}_2} \rangle_V = \langle \boldsymbol{f}^{\boldsymbol{M}} \rangle_V O[V^{-1}]. \tag{4.1.33}$$

(4) The above consideration implies that we may ignore in the large volume limit all the cumulants $\langle \boldsymbol{f}^{\boldsymbol{M}} \rangle_C$ with $\boldsymbol{M}$ having some component(s) larger than 1.

Now, (4.1.31) reads (notice that $\boldsymbol{M}! = 1$)

$$\log Q = \sum_{\boldsymbol{M} \neq 0} \left\langle \boldsymbol{f}^{\boldsymbol{M}} \right\rangle_V . \tag{4.1.34}$$

The summation is over all the choices of 1-PI diagrams with no multiple edges connecting two vertices directly. If $\boldsymbol{f}^{\boldsymbol{M}}$ corresponds to a $k$ vertex 1-PI diagram, there are $\binom{N}{k} \sim N^k$ ways to choose $k$ particles. Because $\langle \boldsymbol{f}^{\boldsymbol{M}} \rangle = O[1/V^{k-1}]$, the overall contribution of such diagrams to the summation is proportional to $N^k/V^{k-1} \sim Nn^{k-1}$. Since $\log Q$ must be extensive, this is just the right contribution from such diagrams:

$$\log Q = \sum_{k=2}^{\infty} \binom{N}{k} \sum_{D(k)} \left\langle \boldsymbol{f}^{\boldsymbol{M}} \right\rangle_{V,k}, \tag{4.1.35}$$

where $D(k)$ is the 1-PI silhouettes (see Fig. 4.1.5) with $k$ vertices:

$$\sum_{D(k)} \left\langle \boldsymbol{f}^{\boldsymbol{M}} \right\rangle_{V,k} = \frac{1}{V^{k-1}} \int d(\text{independent relative particle coordinates}) \sum \boldsymbol{f}^{\boldsymbol{M}}. \tag{4.1.36}$$

The sum here is over all the topologically distinguishable assignments of the particles $1, \cdots, k$ to the vertices of the silhouette $D(k)$. For example, for $k = 4$ there are three distinct silhouettes as shown in Fig. 4.1.5(b), and for one of them there are ways to assign 4 particles as exhibited in (a). If the silhouettes are identical, then the integrated values are identical, so we have only to count the ways for such assignments. For the example (a) there are 6 ways.

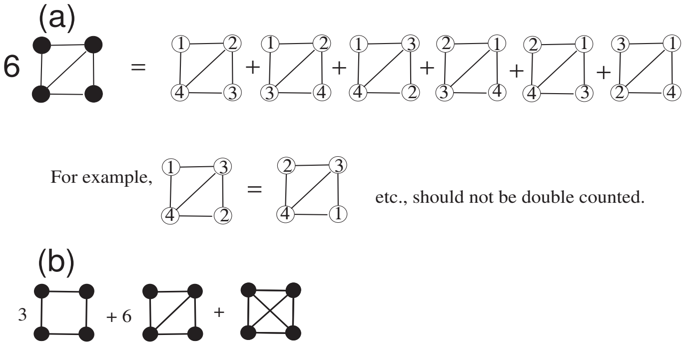


Fig. 4.1.5

(a) "Silhouette' implies the diagrams that are identical if particles are indistinguishable.

(b) For $k = 4$ there are three different kinds of silhouettes (topologically different silhouettes), each of which have several ways (as denoted with a prefactor) of assigning 4 particles.

Finally, we obtain the virial expansion of the free energy:

$$\frac{1}{N}\log Q = \sum_{k=1}^{\infty} \frac{n^k}{k+1}\beta_k, \tag{4.1.37}$$

where

$$\beta_k = \frac{1}{k!}\int d(\text{independent relative particle coordinates})(\text{all 1-PI } k+1 \text{ particle silhouettes}*). \tag{4.1.38}$$

'Silhouette∗' implies that the number of ways counted as in Fig. 4.1.5 has been taken into account (see (4.1.41); implying that, for example, the numerical coefficients as 3 and 6 in this formula should be included). This integral is called the *k-irreducible cluster integral*. For example,

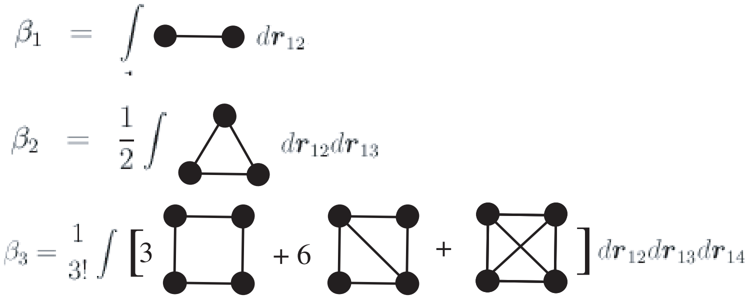
$$\beta_1 = \int f_{12}d\boldsymbol{r}_{12}, \tag{4.1.39}$$

$$\beta_2 = \frac{1}{2}\int f_{12}f_{23}f_{31}d\boldsymbol{r}_{12}d\boldsymbol{r}_{13}. \tag{4.1.40}$$

$\beta_3$ consists of three different silhouettes:

$$\beta_3 = \frac{1}{3!}\int \left(3f_{12}f_{23}f_{34}f_{41} + 6f_{12}f_{23}f_{34}f_{41}f_{13} + f_{12}f_{23}f_{34}f_{41}f_{13}f_{24}\right) d\boldsymbol{r}_{12}d\boldsymbol{r}_{13}d\boldsymbol{r}_{14}. \tag{4.1.41}$$

Diagrammatic expressions are as follows:



From (4.1.37) and (4.1.4) we finally obtain the virial expansion of equation of state:

$$\frac{PV}{Nk_BT} = 1 - \sum_{k=1}^{\infty}\frac{k}{k+1}\beta_k n^k, \tag{4.1.42}$$

$$= 1 + B(T)n + C(T)n^2 + D(T)n^3 + \cdots. \tag{4.1.43}$$

Notice that when we compute the virial coefficients, $V, N \to \infty$ is taken with $n$ being kept constant. The series expansion and this limit are not commutative. Therefore, from the nature of the series (4.1.43) we cannot conclude anything about the existence or absence of phase transition.[11] It is known that this series has a finite convergence radius.[12]

## 4.2 Van der Waals equation of state

Van der Waals[13] proposed the following equation of state (*van der Walls equation of state*:

$$P = \frac{Nk_BT}{V - Nb} - \frac{aN^2}{V^2}, \tag{4.2.1}$$

where $a$ and $b$ are materials constants. Here, $P, N, T, V$ have the usual meaning in the equation of state of gases. His key ideas are:
(1) The existence of the real excluded volume due to the molecular core should reduce the actual volume from $V$ to $V - Nb$; this would modify the ideal gas law to $P_{HC}(V - Nb) = Nk_BT$. Here, subscript $HC$ implies 'hard core.'
(2) The attractive binary interaction reduces the actual pressure from $P_{HC}$ to $P = P_{HC} - a/(V/N)^2$, because the wall-colliding particles are actually pulled back by their fellow particles in the bulk.

The most noteworthy feature of the equation is that liquid and gas phases are described by a single equation. Maxwell was fascinated by the equation, and gave the liquid-gas coexistence condition (*Maxwell's rule*).**q**

---

[11] However, this does not mean that we cannot obtain the critical temperature from the coefficients. See T. Kihara and J. Okutani, Chem. Phys. Lett., **8**, 63 (1971). See Problem 5.1.

[12] ⟪**Virial expansion converges**⟫**r** Theorem [Lebowitz and Penrose] The radius of convergence of the virial expansion (4.1.42) is at least $0.289\cdots(e^{2\beta B}+1)^{-1}C(T)^{-1}$, where $-2B$ is the lower bound of $\phi$. □
See D. Ruelle, *Statistical Mechanics* (World Scientific, 1999) p 85.

[13] ⟪**van der Waals' biography**⟫ See the following page for his scientific biography:
http://www.msa.nl/AMSTEL/www/Vakken/Natuur/htm/nobel/physics-1910-1-bio.htm.

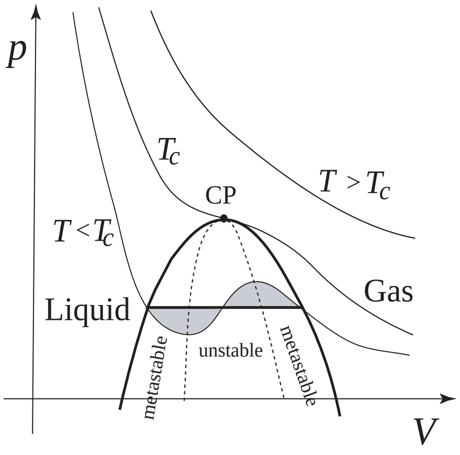


Fig. 4.2.1 The thick curve is the *coexistence curve* below which no single phase can stably exist, and the dotted curve is the *spinodal curve* bounding the thermodynamically unstable states; the region between the spinodal and coexistence curves is the metastable region. When a high temperature state is quenched into the unstable region, it immediately decomposes into liquid and gas phases. If a high temperature state is quenched into the metastable region, after nucleation (of bubbles or droplets) of the minority phase, phase separation occurs. The liquid-gas coexistence pressure for a given temperature is determined by *Maxwell's rule*: the two shaded regions have the same area.

*Maxwell's rule* is motivated by the calculation of $G$: $dG = VdP$. Thus, we can compute the area between the $P$-axis and the $PV$ curve in Fig. 4.2.1. The result is illustrated in Fig. 4.2.2.

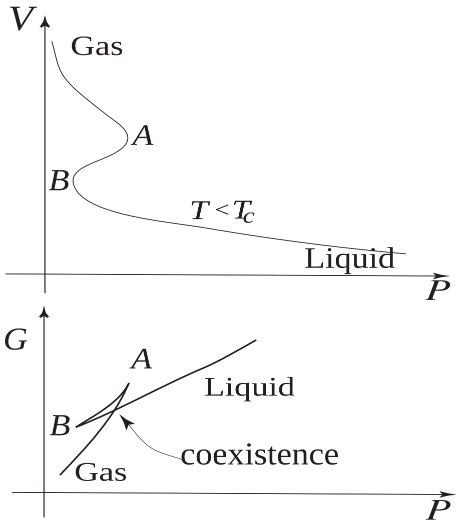


Fig. 4.2.2r Maxwell's rule explained. $G$ is essentially the (signed) area below the $PV$ curve. Thus, $A$ and $B$ give 'spines.' The spinodal curve owes its name to these spines. The self-crossing point of the $G$ curve corresponds to the equal shaded area in Fig. 4.2.1.

However, it is not generally justifiable to use thermodynamics when the system is unstable (definitely no equilibrium is possible), so the above motivation cannot be justified, but as we see in a Problem* to this chapter, Maxwell's rule can be properly justified without abusing thermodynamics.

The van der Waals equation of state is heuristically derived, but what is really the microscopic model that gives it, if any? A proper understanding of van der Waals's idea is[14]

$$P = P_{HC} - \frac{1}{2} a n^2, \tag{4.2.2}$$

where $P_{HC}$ is the hard core fluid pressure, and the subtraction term is the average effect of attractive forces.[15] The equation of state of this type is called (by Widom) the *augmented van der Waals equation of state*.[16] As it is, this equation exhibits the non-monotonic (i.e., not thermodynamically realizable) $PV$ curve just as the van der Waals equation of state, so there cannot be any microscopic model for (4.2.2).[17] However, this equation augmented with Maxwell's rule is thermodynamically legitimate, and indeed it is the equation of state of the gas interacting with the *Kac potential*:**rq**

$$\phi(r) = \phi_{HC}(r/\sigma) + \gamma^3 \phi_0(\gamma r/\sigma), \tag{4.2.3}$$

where $\phi_{HC}(x)$ is the hard core potential: 0 beyond $x = 1$ and $\infty$ for $x \le 1$, and $\phi_0$ is an attractive tail. The parameter $\gamma$ is a scaling factor; we consider the $\gamma \to 0$ limit (long range but infinitesimal interaction); note that the second virial coefficient is independent of $\gamma$.

Theorem.[18] The equation of state for the system with the Kac potential is given by**r**

$$P \equiv \lim_{\gamma \to \infty} p(\gamma) = \left[ \frac{1}{2} n^2 \overline{\phi_0} + P_{HC}(v) \right]_{Maxwell}. \tag{4.2.4}$$

Here, 'Maxwell' implies the application of Maxwell's rule, $P_{HC}$ is the pressure of the hard core fluid (see Section 4.3) and

$$\overline{\phi_0} = \int \phi_0(r) 4\pi r^2 dr. \tag{4.2.5}$$

□

---

[14] Generally speaking, van der Waals's idea of separating the interaction into the hard core part and the long-range attractive part is a productive one. See, for example, D. Chandler, J. D. Weeks and H. C. Andersen, "Van Der Waals Picture of Liquids, Solids, and Phase Transformations," Science **220**, 787-794 (1983).

[15] For a collection of hard cores there is no gas-liquid transition at any temperature.

[16] This is not so impressive for dense fluid or liquid.

[17] Roughly speaking, if the interaction potential is not too long-ranged, if it does not allow pushing infinitely many particles into a finite volume, and if the total intereaction energy is bounded from below, then the normal thermodynamics is obtained.

[18] P. C. Hemmer, M. Kac and G. E. Uhlenbeck, "On the van der Waals theory of the liquid-vapor equilibrium, I.," J. Math. Phys. **4**, 216 (1963). This is for 1D, and is later extended to 3D by J. L. Lebowitz and O. Penrose, "Rigorous treatment of the van der Waals-Maxwell theory of the liquid-vapor transition," J. Math. Phys., **7**, 98-113 (1966).

We can exactly compute the equation of state of a 1D fluid interacting with the Kac potential,r because we can compute $P_{HC}$ exactly in 1-space (hard-rod gas, Tonks' gas).[19] Consider a 1D box of length $L$ containing $N$ particles whose hard core size (length) is $\sigma$. We discretize the space into a lattice of spacing $\delta$. Let $M = L/\delta$, and $m = \sigma/\delta$ (assume both are integers). That is, the discretization converts the problem to the one of placing $N$ rods consisting of $m$ lattice points on the lattice with total $M$ lattice points without any overlap. This problem is equivalent to distributing $M - Nm$ vacant points into $N + 1$ positions (rod spacings + 2 end spaces). This is equivalent to distributing $M - Nm$ points into $N + 1$ bins. Thus, the total number $Z$ of the configurations is given by

$$\binom{M - Nm + N}{N}. \tag{4.2.6}$$

Since the contribution of the kinetic energy to the canonical partition function is irrelevant to the pressure, we have only to compute

$$\frac{P}{k_BT} = \frac{\partial A}{\partial L} = \frac{1}{\delta}\frac{\partial A}{\partial M} = \log\left(\frac{M - Nm + N}{M - Nm}\right) = \frac{1}{\delta}\log\left(1 + \frac{N\delta}{L - N\sigma}\right). \tag{4.2.7}$$

Now, take the $\delta \to 0$ limit to obtain

$$\frac{P}{Nk_BT} = \frac{1}{V - b}, \tag{4.2.8}$$

where $L$ is written as $V$ (volume) and $b = N\sigma$ is the total excluded volume. Therefore, the 1D Kac model gives

$$P = \left[\frac{Nk_BT}{V - b} - \frac{1}{2}a\left(\frac{N}{V}\right)^2\right]_{Maxwell}. \tag{4.2.9}$$

This equation gives without Maxwell's rule thermodynamically unstable states, so a liquid-gas phase transition exists (due to the infinite-range interaction).

We have realized that van der Waals equation of state corresponds to an idealization limit of mean-field like attractive interaction with the hard core to maintain the system stable. It is a theoretically very respectable equation.

---

[19] L. Tonks, "The complete equation of state of one, two and three-dimensional gases of hard elastic spheres," Phys. Rev. **50**, 955 (1936). This work computes the partition function honestly. Actually, we can do better as in a problem* at the end of this chapter (*Takahashi gas*)

van der Waals later realized that the materials constants $a$ and $b$ can be gotten rid of in terms of the critical parameters $T_c, V_c$ and $P_c$ (critical temperature, volume and pressure, respectively;[20] see Problem; critical points will be discussed in Chapter 4): introducing reduced quantities $T_r = T/T_c$, $P_r = P/P_c$ and $V_r = V/V_c$, the equation of state can be written as

$$P_r = \frac{(8/3)T_r}{V_r - 1/3} - \frac{3}{V_r^2}. \tag{4.2.10}$$

He proposed the *law of corresponding states*: in terms of reduced quantities, all the gas equations of state can be expressed by a universal equation of state.[21] Certainly, the gas equation of state is universal in the dilute limit (the ideal gas law); this universality is a trivial one. Near the critical point the equation of state becomes again universal (this is a nontrivial one as we will see later). In other states there is no deep reason to believe the universality claimed by van der Waals. However, if the binary interaction potential is, for example, given by the Lennard-Jones 12-6 potential that has two parameters (energy scale $\varepsilon$ and the length scale $\sigma$), then dimensional analysis tells us that the reduced equation of state must be universal. This is a trivial universality. Unfortunately, the intermolecular forces are not that universal, so the law of corresponding states remains as an approximate law. However, it is practically sufficiently useful, and Kammerlingh-Onnes used it as a guide-line to liquify helium.

## 4.3 Liquid state

The most difficult classical many body problem in equilibrium statistical mechanics is liquid.[22] P. A. Egelstaff, *An Introduction to the Liquid State* (Oxford, 1994) is an excellent introduction.[23] Again, liquid is explained because some methods and ideas are probably useful in wider contexts. For example, the use of functional Taylor

---

[20] In this case the critical temperature is defined by the temperature where the unstable state ceases to exist. The *critical point* is the inflection point of the equation of state at this temperature.

[21] However, for a gas to be described by a van der Waals equation $P_cV_c/Nk_BT_c = 3/8$ is required. Thus, it is obvious that this cannot be universal.

[22] Phase transitions are difficult, but when we study a single phase made of spherically symmetric particles with short-range interaction forces liquid is the most complicated.

[23] J. M. Ziman, *Models of Disorder, the theoretical physics of homogeneously disordered systems* (Cambridge UP, 1979) contains many readable related discussions.

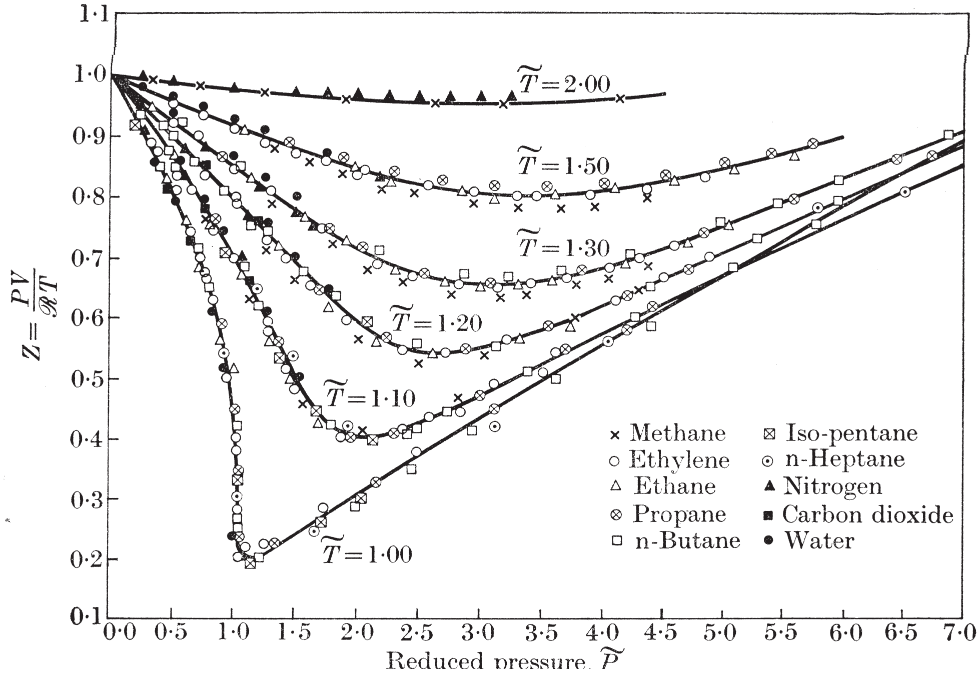


Fig. 4.2.3. The law of corresponding states.

expansion[24] may be of some value to take note.

As we have seen van der Waals could convince the people that liquid and gas are continuously connected without any phase transition. However, near the triple point, solid, liquid, and gas phases are quite distinct. What are the characteristics of these phases?[25] As can be seen from the magnitude of the latent heat between them and their densities solids and liquids are rather similar. We know that solids are ordered, and liquids are not so, but then from this point of view, it is not easy to distinguish liquid from gas. The characterization by Bernal, who is the founder

---

[24] A function of functions, that is, a map that maps functions to something else (e.g., numbers, functions) is called a functional.

[25] ⟪**Phases**⟫$_{\mathbf{r}}$ It is not easy to define phases globally in the phase diagram as can be seen from the example of gas and liquid phases (see also Chapter 5). Therefore, in this book, the concept of 'phase' is used 'locally' when precise statements are needed. We say the states (near the phase transition point) that cannot be changed into each other without a phase transition (= thermodynamic singularity, see Section 5.3) are distinct phases (near the phase transition point).

of molecular biology,[26] seems to be the best:

| | long range order | coherence |
|---|---|---|
| solid | ○ | ○ |
| liquid | × | ○ |
| gas | × | × |

Here, *coherence*$_{\mathbf{r}}$ means that at any instant at least four particles are repulsively interacting with a chosen particle, and such repulsive contacts structurally resist against compression. This is the microscopic reason why liquid cannot be compressed so easily as gas. Thus, it is clear that repulsive interactions due to hard cores of molecules are crucial. Therefore, we may say that the *ideal liquid* is the dense hard core fluid confined in a box (the box replaces the collective cohesive effect due to attractive intermolecular interactions; recall the Kac potential and augmented van der Waals equation of state).

The lack of long range order but the existence of coherence imposes a peculiar local structure (approximate symmetry) to the core packing in liquid. If there is a 3, 4 or 6 fold symmetry around a given particle, this local symmetry is compatible with a long range order as we know from the crystal structure such as face-centered cubic lattice, cubic lattice, or hexagonal lattice. Bernal realized that local five-fold symmetry is the key to the liquid molecular packing. Studying the actual ball-bearing 'random close packing',[27] he found the so-called Bernal polyhedra.[28] They represent the polyhedra whose vertices correspond to the centers of the particles touching a given particle in liquid.[29]

Because of this local packing flexibility the mechanism of thermal expansion of liquid is distinct from solid. In solids, the volume expands by increasing the interparticle distance without modifying the structure. However, in liquids, particles can be rearranged locally, so the volume can increase by reducing the coordination number (increasing vacancies; however, do not imagine any vacancy or void that can accommodate a particle). Thus generally, the expansion coefficient of liquid is larger

---

[26] See a recent book review of his biography in Nature, **440**, 149 (2006).

[27] ⟪**Random close packing well defined?**⟫ The requirement of randomness and that of closeness are contradictory, so there is a criticism that such a structure is not well-defined and is not worthy of any precise discussion. However, here Bernal's intuition is trusted and the viewpoint is adopted that we cannot precisely characterize such a structure, not because it is intrinsically ill-defined, but because we have not yet been able to characterize the structure of liquid.

[28] J. D. Bernal, "Geometry of the structure of monatomic liquids," Nature **185**, 68 (1960).

[29] There is a work observing the five-fold symmetry: H. Reichert, O. Klein, H. Dosch, M. Denk, V. Honkimäki, T. Lippmann and G. Reiterk, "Observation of five-fold local symmetry in liquid lead," Nature **408**, 839 (2000).

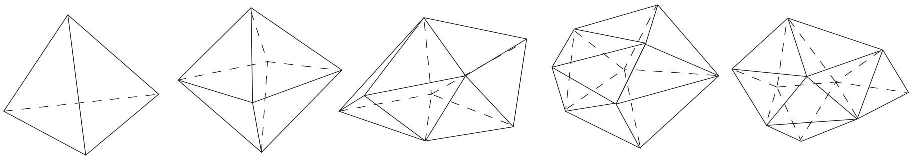


Fig. 4.3.1 Bernal polyhedra representing the local core packing. Notice the local five-fold symmetry in three of them.

than the corresponding solid.[30]

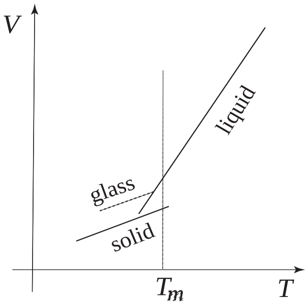


Fig. 4.3.2
Thermal expansion of liquid and solid. As long as the particles can be rearranged locally, the expansion coefficient (slope) is large even in supercooled liquid. However, upon dynamical freezing (glass formation) the expansion mechanism becomes just as in solids. Here, $T_m$ is the melting point.

To study the structure more quantitatively, it is convenient to count the number of particles located between $r$ and $r + dr$ from (the center of) a particular molecule as

$$4\pi n g(r) r^2 dr = \text{the number of particle ‘centers’ in the shell between } r \text{ and } r + dr. \tag{4.3.1}$$

Here, $n$ is the number density, and $g(r)$ is called the *radial distribution function* (see (4.3.10) analytically). The general shape of the radial distribution function is given in Fig. 4.3.3. The number of particles under the first peak is the *coordination number*. This is about 8 for argon at the triple point.

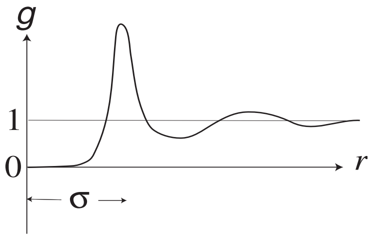


Fig. 4.3.3
A typical shape of the radial distribution function $g$. $\sigma$ is roughly the diameter of the particles making the liquid. The first peak corresponds to the first coordination shell, and the second peak to the second coordination shell (see Fig. 4.3.4).

---

[30] Structural liquids such as water can be different.

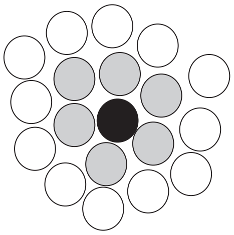


Fig. 4.3.4
The gray particles make the first coordination shell, and the surrounding particles make the second coordination shell, roughly.

The radial distribution function is the quantity of central importance for liquids:
(i) It is directly observable by (e.g., neutron) scattering experiments.
(ii) It is related to the potential of mean force between two particles.
(iii) If the interaction is binary, it describes all the thermodynamic quantities of liquid. (This is a good exercise, so is left to the reader.[31])

Before explaining (i)-(iii), let us summarize many-body distribution functions. Instantaneous $k$-body (density) distribution function is defined as

$$\rho^{(k)}(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_k)=\sum_{i_1,\dots,i_k\in\{1,\cdots,N\}}\delta(\boldsymbol{r}_1-\boldsymbol{x}_{i_1})\cdots\delta(\boldsymbol{r}_k-\boldsymbol{x}_{i_k}), \tag{4.3.2}$$

where $i_1,\cdots,i_k$ are all distinct, $\boldsymbol{x}_i$ is the position vector of the $i$-th particle. If this is averaged over an equilibrium distribution, the $k$-body (density) distribution function is obtained:

$$n^{(k)}(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_k)=(N)_k\langle\delta(\boldsymbol{r}_1-\boldsymbol{x}_1)\cdots\delta(\boldsymbol{r}_k-\boldsymbol{x}_k)\rangle, \tag{4.3.3}$$

where $(N)_k=N(N-1)\cdots(N-k+1)$ (a standard notation). In particular, the single-body distribution function is defined as

$$n^{(1)}(\boldsymbol{x})=\langle\rho^{(1)}(\boldsymbol{x})\rangle \tag{4.3.4}$$

in terms of the instantaneous single body distribution function

$$\rho^{(1)}(\boldsymbol{r})=\sum_{i\in\{1,\cdots,N\}}\delta(\boldsymbol{r}-\boldsymbol{x}_i). \tag{4.3.5}$$

Analogously, the two-body distribution function is defined as

$$n^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)=\langle\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\rangle, \tag{4.3.6}$$

---

[31] The reader must know $g$ at various temperatures, and integration with respect to temperature may be required. The answer may be found in Egelstaff's book.

s

where $\rho^{(2)}$ is the instantaneous two-body distribution function defined as

$$\rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \sum_{i_1, i_2 \in \{1, \cdots, N\}, i_1 \neq i_2} \delta(\boldsymbol{r}_1 - \boldsymbol{x}_{i_1})\delta(\boldsymbol{r}_2 - \boldsymbol{x}_{i_2}). \tag{4.3.7}$$

Here, the sum is over all the way ($N(N-1)$ ways) to choose distinct ordered pairs of particles from $N$ distinguishable particles. If we know the $k$-body distribution function, we can obtain the $(k-1)$-body distribution functions by integration as

$$\int_V d^3\boldsymbol{r}_k \, n^{(k)}(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_k) = (N-k+1)n^{(k-1)}(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_{k-1}). \tag{4.3.8}$$

In particular,

$$\int_V d^3\boldsymbol{r} \, n^{(1)}(\boldsymbol{r}) = N, \tag{4.3.9}$$

so $n^{(1)}$ behaves roughly as the number density $n$ (if the system is uniform, it is the number density itself: $n^{(1)}(\boldsymbol{x}) = N\langle\delta(\boldsymbol{x} - \boldsymbol{r}_1)\rangle = N/V = n$). We see that $n^{(2)}$ behaves like $n^2$. If the system is isotropic and uniform, we define the radial distribution function $g$ as

$$n^{(2)}(\boldsymbol{r}, \boldsymbol{r}') = n^2 g(|\boldsymbol{r} - \boldsymbol{r}'|). \tag{4.3.10}$$

$g(r)$ is zero for small $r$ due to the repulsive core. In the $r \to \infty$ limit, the particle-particle correlation should disappear, so $n^{(2)}(\boldsymbol{r}, \boldsymbol{r}') \to n^2$, implying $g \to 1$.

The scattering amplitude $A$ of, e.g., neutron monochromatic beam from the instantaneous density distribution $\rho(\boldsymbol{x})$ may be written as

$$A \propto \int e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \rho(\boldsymbol{x}) d\boldsymbol{x}, \tag{4.3.11}$$

so the scattering function is proportional to

$$S(\boldsymbol{k}) = \langle |A|^2 \rangle \propto \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j)} \right\rangle. \tag{4.3.12}$$

The term due to $i = j$ does not depend on $\boldsymbol{k}$, so we may ignore it. Now, the radial distribution function may be written as (see (4.3.6) for $n^{(2)}$)[32]

$$n^2 g(\boldsymbol{r}) = n^{(2)}(\boldsymbol{r}, 0) = \langle \rho(\boldsymbol{r})\rho(0) \rangle. \tag{4.3.13}$$

---

[32] The rightmost term in the following formula is $\rho^{(2)}(\boldsymbol{r}, \boldsymbol{r}') = \rho^{(1)}(\boldsymbol{r})\rho^{(1)}(\boldsymbol{r}') + N\delta(\boldsymbol{r} - \boldsymbol{r}')$, so the equality is not right, but the difference is where there is no $\boldsymbol{k}$ dependence (or $\boldsymbol{k} = 0$ term, that is the forward scattering term), so it is ignored.

Thus, the scattering function (form factor) $S(\boldsymbol{k})$ is proportional to the Fourier transform of $g$.[33]

Let $w(\boldsymbol{r})$ be the required reversible work to bring one particle from infinity (outside the fluid) to $\boldsymbol{r}$, when another particle is fixed at the origin in the fluid. Then,

$$g(\boldsymbol{r}) = \exp(-\beta w(\boldsymbol{r})). \tag{4.3.14}$$

$w$ is called the *potential of mean force*.**rq**

Let us write the total potential energy to be $U(\{\boldsymbol{r}_i\})$, and $\boldsymbol{r}_1 = \boldsymbol{r}$, $\boldsymbol{r}_2 = 0$. The mean force required to move the first particle is

$$\begin{aligned} -\frac{d}{d\boldsymbol{r}}\langle U\rangle_{\boldsymbol{x}_1=\boldsymbol{r},\boldsymbol{x}_2=0} &= -\frac{\int d\boldsymbol{x}_3\cdots\boldsymbol{x}_N(\partial U/\partial\boldsymbol{r})e^{-\beta U}}{\int d\boldsymbol{x}_3\cdots\boldsymbol{x}_N e^{-\beta U}} && (4.3.15)\\ &= k_BT\frac{d}{d\boldsymbol{r}}\log\int d\boldsymbol{x}_3\cdots\boldsymbol{x}_N e^{-\beta U}, && (4.3.16)\\ &= k_BT\frac{d}{d\boldsymbol{r}}\log g(\boldsymbol{r}), && (4.3.17)\end{aligned}$$

because the normalization constant does not depend on the particle positions. Integrating this, we conclude that $w$ is the reversible work required to bring the first particle to $\boldsymbol{r}$ while fixing the second particle at the origin.

## 4.4 Statistical mechanics of simple liquid

As we have seen in the preceding section, to obtain $g$ is the central issue of liquid state, especially for simple liquids consisting of spherical particles. To this end it is convenient to analyze the structure of $g$. $g \to 1$ in the long distance limit, so

$$h(\boldsymbol{r}) = g(\boldsymbol{r}) - 1 \tag{4.4.1}$$

is a more convenient function, which is called the *indirect correlation function.* The damping oscillatory nature of $h$ may be intuitively understood from the successive coordination shells as in Fig. 4.3.4.

Ornstein and Zernike thought the oscillatory behavior of $h$ may be constructed

---

[33] In some cases (esp., with X-ray scattering) there is a $\boldsymbol{k}$-dependent prefactor due to a particular scattering mechanism, so its effect must be eliminated from $S(\boldsymbol{k})$ to obtain $g$.

from a short range function (a function with a small support near the origin) $c$ successively as

$$h = c + nc * c + n^2 c * c * c + \cdots , \tag{4.4.2}$$

where $*$ is the spatial convolution:

$$c * c(\boldsymbol{r}) = \int d\boldsymbol{r}'\, c(\boldsymbol{r} - \boldsymbol{r}')c(\boldsymbol{r}'). \tag{4.4.3}$$

Summing over this sequence, we may write

$$h(\boldsymbol{r}) = c(\boldsymbol{r}) + n \int d\boldsymbol{r}'\, c(\boldsymbol{r}')h(\boldsymbol{r} - \boldsymbol{r}'). \tag{4.4.4}$$

Mathematically, $c$ is defined by this equation called the *Ornstein-Zernike equation*. $c$ is called the *direct correlation function* and is indeed shorter ranged than $h$ (at least away from the critical point if the binary interactions are short-ranged). If we can obtain one more relation between $h$ (or $g$) and $c$ that involves $\phi$, we can obtain $h$.

The strategy that may be recommended to obtain this second relation is based on the following series of ideas:
(1) $g$ is obtained from the two body correlation function $n^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)$. This can be obtained if we can compute the single body distribution $n^{(1)}(\boldsymbol{r}_1|U)$ under the existence of an arbitrary external potential $U$. Set $U(\boldsymbol{r}) = \phi(\boldsymbol{r} - \boldsymbol{r}_2)$; actually, $n^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = nn^{(1)}(\boldsymbol{r}_1|\phi)$. This is, physically, equivalent to pushing in the second particle at $\boldsymbol{r}_2$ (the idea may be realized numerically[34]).
(2) Therefore, to study $n^{(1)}(\boldsymbol{r}_1|U)$ is the key. To this end we study the effect on $n^{(1)}(\boldsymbol{r}_1|U)$ of perturbation of $U$:

$$\delta n^{(1)}(\boldsymbol{r}|U) = Q\delta U. \tag{4.4.5}$$

If we know the linear operator $Q$,[35] then, since it is a sort of derivative, 'functional version' of the fundamental theorem of calculus should allow the 'integration' to obtain $n^{(1)}(\boldsymbol{r}_1|U)$.
(3) Unfortunately, we cannot obtain $Q$ for general $U$, so this functional integral program cannot be performed exactly. We use a 'linear approximation' with $Q$ evaluated at $U = 0$.

---

[34] By a Monte Carlo method, see R. Pomès, E. Eisenmesser, C. B. Post and B. Roux, "Calculating excess chemical potentials using dynamic simulations in the fourth dimension," J. Chem. Phys., **111**, 3387 (1999); see also a recent paper by Pomès et al., "Absolute free energy calculations by thermodynamic integration in four spatial dimensions," J. Chem. Phys., **123**, 034104 (2005).

[35] Mathematically, this is the Frechét derivative.

(4) We wish to know an approximate relation between $n$ and $U$. The direct application of the above idea may not be the best. A natural idea is to improve an approximate relation $A \simeq B$ by linear Taylor approximation, where $A$ and $B$ are functionals of $n$ and $U$. (If we regard the procedure as an improvement of $A(U) - A(0) \simeq B(U) - B(0)$, then $A$ and $B$ need not be close at all).[36] If teh following exposition is preempted, $Q = (\delta A/\delta B)_{U=0}$ and formally, what we wish to write down as an improved equation is, formally,$_{\mathbf{q}}$

$$A(U) - A(0) = \left.\frac{\delta A}{\delta B}\right|_{U=0} [B(U) - B(0)]. \tag{4.4.6}$$

To implement the above program, we must be able to compute something (= functional derivative) corresponding to $Q$ in (4.4.5) for $A$, $B$. A brief summary of the basic calculation rules of functional derivatives is given below.

A map $F$ that maps a function $g$ to another function $f$ is called a functional:

$$f = F[g]. \tag{4.4.7}$$

Here, the value of $f$ at $x$ may be determined by the function $g$ as a whole (i.e., $f(x)$ is not determined by the value of $g$ at $x$ alone). $F$ is called a *functional*. For example, the map from $f(x)$ to its Fourier transform $\hat{f}(x)$ is a functional. This relation may be understood as a relation between the sequence of values $\{g(x)\}$ and that of $\{f(x)\}$. *Functional analysis regards a function $f$ as a vector whose components are $\{f(x)\}$* or $f = (\text{——}f(x)\text{——})$. Let us recall the ordinary vector analysis. If $F$ maps a vector $\boldsymbol{u}$ to $\boldsymbol{v}$ as $\boldsymbol{v} = F(\boldsymbol{u})$, and if $F$ is differentiable, we may study the effect of small change in $\boldsymbol{u}$ on $\boldsymbol{v}$ as

$$v_i + \delta v_i = F_i(\boldsymbol{u}) + \sum_j DF(\boldsymbol{u})_{ij} \delta u_j + \cdots \tag{4.4.8}$$

Analogously, if we change $g$ slightly in (4.4.7), we expect $f$ to change slightly,[37] and we may write

$$f(x) + \delta f(x) = F[g](x) + \int dy DF[g](x,y) \delta g(y) + \cdots, \tag{4.4.9}$$

---

[36] For example, an 'approximate relation' $\cos x \simeq 1 + x^2$ near $x = 0$ does not give a correct relation between the slopes, but if this is 'improved' by the linear Taylor approximation, we obtain $\cos x = 1 - \frac{1}{2}x^2$, which is correct to $O[x^2]$.

[37] As the reader expect from the elementary calculus, just as in the case of the ordinary functions, we need many technical conditions for a functional to be 'differentiable,' but we proceed only formally. No mathematical details such as 'Banach space' are required by our formal applications. Let's be blissfully formal.

where $DF[g](x, y)$ is 'the $(x, y)$ component' of a linear operator analogous to the matrix component $DF(\boldsymbol{u})_{ij}$. The following correspondence explains the intuitive meaning of functional derivative $DF[g](x, y) = \delta f(x)/\delta g(y)$:

$$\delta v_i = \sum_j \frac{\partial v_i}{\partial u_j}\delta u_j \;\leftrightarrow\; \delta f(x) = \int dy\, \frac{\delta f(x)}{\delta g(y)}\delta g(y). \tag{4.4.10}$$

As we will see below, many functional derivatives may be computed fairly easily with the aid of rather mechanical rules, but if some difficulty is encountered, go back to the variational formula (4.4.9). This definition gives us

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x - y). \tag{4.4.11}$$

This should be compared with

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}, \tag{4.4.12}$$

which is the derivative of one component with respect to another of a finite dimensional vector $\boldsymbol{x} = (x_1, \cdots, x_D)$; $\delta_{ij}$ is obtained, because $x_i$ and $x_j$ are independent variables for $i \neq j$. As already noted above, "functional calculus is the calculus regarding a function $f$ as a vector (——$f(x)$——)." Suffixes are not discrete $i = 1$, $2, \cdots$ but continuous $x$. It may not be hard to grasp the general computational rules from a few examples: If $F$ is an ordinary function,

$$\frac{\delta F(f(x))}{\delta f(y)} = F'(f(x))\delta(x - y) \tag{4.4.13}$$

is a chain rule. The next example also illustrates a chain rule, when $G$ is a functional:

$$\frac{\delta G[f(x)]}{\delta g(y)} = \int dz\, \frac{\delta G[f(x)]}{\delta f(z)}\frac{\delta f(z)}{\delta g(y)}. \tag{4.4.14}$$

This should not be hard to see if the reader knows the chain rule in vector analysis. Here, $x$ is often a dummy variable, because $G$ depends on $f$ as a whole (not on its particular value at $x$). The ordinary calculus operations are mere addition or subtraction, so we may freely change the order of functional differentiation and the ordinary calculus operations:

$$\frac{\delta f'(x)}{\delta f(y)} = \frac{d}{dx}\delta(x - y) \tag{4.4.15}$$

$$\frac{\delta}{\delta f(x)}\int_A f(y)dy = \int_A \delta(x - y)dy = \chi_A(x). \tag{4.4.16}$$

where $\chi_A$ is the indicator of the set $A$. Let $\hat{f}$ be the Fourier transform of $f$:

$$\hat{f}(k) = \int dx\, f(x) e^{ikx}. \tag{4.4.17}$$

Then,

$$\frac{\delta \hat{f}(k)}{\delta f(x)} = e^{ikx}. \tag{4.4.18}$$

Let $q$ be a twice differentiable function of $t$, and the action is defined as

$$S = \int_{t_1}^{t_2} dt \left[ \frac{1}{2} m \dot{q}^2 - V(q) \right]. \tag{4.4.19}$$

Then, the Newton's equation of motion is just[38]

$$\frac{\delta S}{\delta q(t)} = 0. \tag{4.4.20}$$

Now, we can differentiate the grand partition function with the external potential $U$ to perturb the system. Let us write

$$\Xi[U] = \sum_N \frac{1}{N!} \int d\Gamma_N\, e^{\beta\left[\mu N - H_N - \sum_i U(\boldsymbol{x}_i)\right]}. \tag{4.4.21}$$

Here, $d\Gamma_N$ is the volume element of the $N$-particle phase space, and $H_N$ is the $N$-particle Hamiltonian. We can compute (all the details are given here, so that the reader may practice calculation)

$$\begin{aligned}
\frac{\delta \Xi[U]}{\delta e^{-\beta U(\boldsymbol{r})}} &= e^{\beta U(\boldsymbol{r})} \frac{\delta \Xi[U]}{\delta(-\beta U(\boldsymbol{r}))} && (4.4.22)\\
&= e^{\beta U(\boldsymbol{r})} \sum_N \frac{1}{N!} \int d\Gamma_N\, \frac{\delta}{\delta(-\beta U(\boldsymbol{r}))} e^{\beta\left[\mu N - H_N - \sum_i U(\boldsymbol{x}_i)\right]} && (4.4.23)\\
&= e^{\beta U(\boldsymbol{r})} \sum_N \frac{1}{N!} \int d\Gamma_N\, e^{\beta\left[\mu N - H_N - \sum_i U(\boldsymbol{x}_i)\right]} \left[ \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{x}_i) \right]
\end{aligned}$$

---

[38] The derivative of $\delta$-function is interpreted just as in the standard distribution theoretical sense:

$$\int \delta'(x) f(x) dx = -\int \delta(x) f'(x) dx = -f'(0).$$

$$\tag{4.4.24}$$

$$= e^{\beta U(\boldsymbol{r})}\Xi[U]\left\langle \sum_{i=1}^{N}\delta(\boldsymbol{r}-\boldsymbol{x}_i)\right\rangle_U . \tag{4.4.25}$$

Here, $\langle\ \rangle_U$ denotes the average over the grand canonical ensemble with the external potential $U$. Using the definition of the single body distribution function (4.3.4), we obtain

$$\frac{\delta \log \Xi[U]}{\delta e^{-\beta U(\boldsymbol{r})}} = e^{\beta U(\boldsymbol{r})} n^{(1)}(\boldsymbol{r}|U), \tag{4.4.26}$$

where $n^{(1)}(\boldsymbol{r}|U)$ is the one-particle distribution with the external field $U$. We need derivative of $n$ wrt $U$. Differentiating (4.4.26) again, we obtain

$$\frac{\delta^2 \log \Xi[U]}{\delta e^{-\beta U(\boldsymbol{r})}\delta e^{-\beta U(\boldsymbol{r}')}} = \frac{\delta}{\delta e^{-\beta U(\boldsymbol{r}')}} e^{\beta U(\boldsymbol{r})}\frac{1}{\Xi[U]}\sum_N \frac{1}{N!}\int d\Gamma_N\, e^{\beta\left[\mu N - H_N - \sum_i U(\boldsymbol{x}_i)\right]}\left[\sum_{i=1}^{N}\delta(\boldsymbol{r}-\boldsymbol{x}_i)\right] \tag{4.4.27}$$

$$= e^{\beta U(\boldsymbol{r})}e^{\beta U(\boldsymbol{r}')}\left[n^{(2)}(\boldsymbol{r},\boldsymbol{r}'|U) - n^{(1)}(\boldsymbol{r}|U)n^{(1)}(\boldsymbol{r}'|U)\right]. \tag{4.4.28}$$

In the preceding calculation do not forget the differentiation of the overall multiplicative factor $e^{\beta U(\boldsymbol{x})}$ (if forgotten, $n^{(2)}$ cannot be obtained). Combining (4.4.26) and (4.4.28), we obtain

$$-\frac{\delta n^{(1)}(\boldsymbol{r}|U)}{\delta \beta U(\boldsymbol{r}')} = e^{-\beta U(\boldsymbol{r}')}\frac{\delta e^{\beta U(\boldsymbol{r})} n^{(1)}(\boldsymbol{r}|U)e^{-\beta U(\boldsymbol{r})}}{\delta e^{-\beta U(\boldsymbol{r}')}} \tag{4.4.29}$$

$$= n^{(2)}(\boldsymbol{r},\boldsymbol{r}'|U) - n^{(1)}(\boldsymbol{r}|U)n^{(1)}(\boldsymbol{r}'|U) + n^{(1)}(\boldsymbol{r}|U)\delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{4.4.30}$$

Therefore,

$$-\left(\frac{\delta n^{(1)}(\boldsymbol{r}|U)}{\delta \beta U(\boldsymbol{r}')}\right)_{U=0} = n^{(2)}(\boldsymbol{r},\boldsymbol{r}') - n^2 + n\delta(\boldsymbol{r}-\boldsymbol{r}') = n[nh(\boldsymbol{r}-\boldsymbol{r}') + \delta(\boldsymbol{r}-\boldsymbol{r}')]. \tag{4.4.31}$$

Here, $h$ is the indirect correlation function .

Let us return to our project with the above preparatory steps. Our strategy is to find an appropriate function(al)s $A$ and $B$ of $n^{(1)}$ and $U$, and improve the relation $\Delta A \simeq \Delta B$ by the Taylor expansion trick. This trick even improves relation $A \not\simeq B$ as already noted. In (4.4.5) $A = n$ and $B = U$. However, $U$ is eventually set to be equal to $\phi$ containing a hard core. Thus, the perturbation cannot be weak at all, and the first order approximation should not be very reliable.[39] $B = f = e^{-\beta U} - 1$ should be far better but 1 is not a very small quantity. In the dilute limit $n \sim e^{-\beta U}$, so $A = e^{\beta U} n$ may be a better choice than $A = n$. Then we could choose $B = n$. Perhaps we could use logarithms of these choices as well. The choice $A = e^{\beta U} n$ and $B = n$ gives the *Percus-Yevick* integral equation theory of liquid, perhaps the 'best' theory that gives realistic $g$.

We wish to Taylor-expand $A = e^{\beta U} n$ with respect to $n$ around $U = 0$. We need

$$\left.\frac{\delta n^{(1)}(\boldsymbol{r}|U)e^{\beta U(\boldsymbol{r})}}{\delta n^{(1)}(\boldsymbol{r}'|U)}\right|_{U=0} = \delta(\boldsymbol{r}-\boldsymbol{r}') + n^{(1)}(\boldsymbol{r}|U)\left.\frac{\delta e^{\beta U(\boldsymbol{r})}}{\delta n^{(1)}(\boldsymbol{r}'|U)}\right|_{U=0}, \tag{4.4.32}$$

$$= \delta(\boldsymbol{r}-\boldsymbol{r}') + n\left.\frac{\delta \beta U(\boldsymbol{r})}{\delta n^{(1)}(\boldsymbol{r}'|U)}\right|_{U=0}. \tag{4.4.33}$$

The derivative is the inverse of (4.4.31) as seen from the following chain rule:

$$\int dy\, \frac{\delta f(x)}{\delta g(y)}\frac{\delta g(y)}{\delta f(z)} = \delta(x-z). \tag{4.4.34}$$

We use the Ornstein-Zernike equation (4.4.4) to compute the inverse of (4.4.31): ($\delta$-function is the unit element for the $*$-product)

$$h = c + nc * h = c * (\delta + nh). \tag{4.4.35}$$

Therefore,

$$nc * (\delta + nh) = nh + \delta - \delta \Rightarrow \delta = (\delta + nh) * (\delta - nc). \tag{4.4.36}$$

This implies that the inverse of $\delta + nh$ is $\delta - nc$, so (4.4.33) reads

$$\left.\frac{\delta n^{(1)}(\boldsymbol{r}|U)e^{\beta U(\boldsymbol{r})}}{\delta n^{(1)}(\boldsymbol{r}'|U)}\right|_{U=0} = nc(\boldsymbol{r}-\boldsymbol{r}'). \tag{4.4.37}$$

The linearized Taylor approximation

$$A(\phi) - A(0) = \int dx \left(\frac{\delta A}{\delta B}\right)_0 [B(\phi) - B(0)] \tag{4.4.38}$$

reads in the present case as

$$ng(\boldsymbol{r})e^{\beta\phi(\boldsymbol{r})} - n = n\int nc(\boldsymbol{r}-\boldsymbol{r}')h(\boldsymbol{r}')d\boldsymbol{r}' \tag{4.4.39}$$

---

[39] However, the Debye-Hückel theory of electrolyte solutions uses this result. See Problem*.

or

$$g(\boldsymbol{r})e^{\beta\phi(\boldsymbol{r})} = g(\boldsymbol{r}) - c(\boldsymbol{r}). \tag{4.4.40}$$

This is called the *Percus-Yevick closure* (PY closure). This has a remarkable property that the support of $c$ is contained in the support of $\phi$. Therefore, for the hard core potential $c$ is nonzero only within the hardcore diameter. The Ornstein-Zernike equation + this closure is the Percus-Yevick theory, and can be solved analytically for spherical hard core fluids in 3D (the *Thiele-Wertheim solution*[40]). The PY equation gives a fairly good approximate description of the radial distribution function.

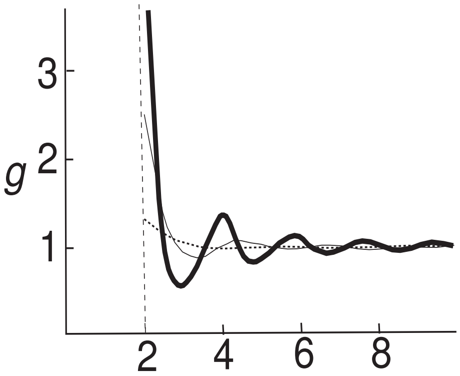


Fig. 4.4.1
A sketch of PY solutions for hard core fluid; the packing density is 0.5 (thick curve), 0.3 (thin curve) and 0.1 (dotted curve). The radius of the hard sphere is 1 (i.e., $\sigma = 2$).

There are two ways to obtain the pressure (i.e., the equation of state) based on $g$; one is based on the *virial theorem* and the other on the *compressibility equation*. These two methods should give the identical equations of state, but our $g$ obtained theoretically is only approximate, so they give different results. The $g$ obtained by the PY closure gives for the hard core fluid (PY-V stands for the equation of state based on the virial theorem, and PY-C that on the compressibility equation)**q**

$$\frac{P}{nk_BT} = \frac{1+2\eta+3\eta^3}{(1-\eta)^2} \text{ PY-V}, \tag{4.4.41}$$

$$\frac{P}{nk_BT} = \frac{1+\eta+\eta^2}{(1-\eta)^3} \text{ PY-C}, \tag{4.4.42}$$

where $\eta = \pi\sigma^3 n/6$ is the *packing fraction*, $\sigma$ is the diameter of the particle and $n$ is the number density. **r**The Monte-Carlo calculation can be fitted very well to the

---

[40] ⟪**Exact solution of the Percus-Yevick theory**⟫ M. S. Wertheim, J. Math. Phys., **5**, 643 (1964). This is a very nice exercise of complex analysis. Actually, it can be solved analytically in any odd dimensional space: E. Leutheusser, "Exact solution of the Percus-Yevick equation for a hard-core fluid in odd dimensions," Physica A **127**, 667 (1984). It is known that the Thiele-Wertheim method allows to solve the Percus-Yevick theory for fluids with any piecewise constant binary potential. ref

following *Carnahan-Stirling equation*:[41]

$$\frac{P}{nk_BT} = \frac{1+\eta+\eta^2-\eta^3/3}{(1-\eta)^3}. \tag{4.4.43}$$

This is between PY-V and PY-C: PY-C > CS > PY-V.

**Virial equation of state**

Clausius derived the *virial theorem*:**rq**

$$\langle K\rangle = -\frac{1}{2}\left\langle \sum q_iF_i \right\rangle, \tag{4.4.44}$$

where $K$ is the total kinetic energy, $F_i$ is the total force acting on the $i$-th particle, and $\langle\ \rangle$ is originally the time average. We know that time average is meaningless in statistical mechanics, so for (4.4.44) to be meaningful macroscopically, the average must be over the canonical distribution.

Notice first that with the aid of Liouville's theorem (invariance of the phase volume)

$$\frac{1}{Z}\int \left(\prod dq(0)dp(0)\right)\sum_i q_i(t)p_i(t)e^{-\beta H} = \frac{1}{Z}\int \left(\prod dq(t)dp(t)\right)\sum_i q_i(t)p_i(t)e^{-\beta H} \tag{4.4.45}$$

$$= \frac{1}{Z}\int \left(\prod dqdp\right)\sum_i q_ip_ie^{-\beta H}, \tag{4.4.46}$$

so

$$\frac{d}{dt}\frac{1}{Z}\int \left(\prod dq(0)dp(0)\right)\sum_i q_i(t)p_i(t)e^{-\beta H} = \frac{1}{Z}\int \left(\prod dq(0)dp(0)\right)\sum_i [\dot{q}_i(t)p_i(t)+q_i(t)\dot{p}_i(t)]e^{-\beta H} = 0. \tag{4.4.47}$$

That is,

$$\left\langle \sum(\dot{p}_iq_i+p_i\dot{q}_i)\right\rangle = 0. \tag{4.4.48}$$

Since $\sum \dot{q}_ip_i = 2K$, this leads to (4.4.44). Its quantum version can be derived in an almost parallel fashion with the aid of Heisenberg's equation of motion.

Let us consider a gas confined in a box made of a potential well $U_W$. Then, the Hamiltonian of the system may be written as $H+U_W$, where $H=K+U$ is the Hamiltonian of the gas system alone (the intrinsic Hamiltonian). (4.4.44) gives us

$$2\langle K\rangle = \sum\left(\left\langle q_i\frac{\partial U}{\partial q_i}\right\rangle + \left\langle q_i\frac{\partial U_W}{\partial q_i}\right\rangle\right). \tag{4.4.49}$$

---

[41] This is an empirical result, obtained by choosing the functional form close to the analytic results.

Since $U_W$ is the effect of the box walls, its derivative is nonzero only very close to the walls. Here, $-\partial U_W/\partial q_i$ is the force the walls exert on the particle. Therefore, if we choose the surface element $d\boldsymbol{S}$ (outward normal) at $\boldsymbol{q}$ on the wall, the sum of $-q_i \partial U_W/\partial q_i$ near the surface element must be $-P\boldsymbol{q} \cdot d\boldsymbol{S}$, where $P$ is the pressure. Notice that $d\boldsymbol{S}$ points outward (outward normal), so $-$ is needed.

$$\sum \left\langle q_i \frac{\partial U_W}{\partial q_i} \right\rangle = \int_{\partial V} \boldsymbol{q} \cdot P d\boldsymbol{S} = 3PV, \tag{4.4.50}$$

where Gauss's theorem has been used. Consequently, (4.4.49) reads**r**

$$PV = \frac{2}{3}\langle K \rangle - \frac{1}{3} \sum \left\langle q_i \frac{\partial U}{\partial q_i} \right\rangle. \tag{4.4.51}$$

This is called the *virial equation of state*. Notice that the general result for the ideal gas $PV = 2E/3$ follows.

**Compressibility in terms of $h$**

We may use functional differentiation to compute $(\partial n/\partial P)_{V,T}$. Let us work with the grand canonical ensemble. In the volume $V$ we apply an infinitesimal uniform external field $\delta U$, and study how $n$ and $P$ vary.[42] Using (4.4.31), we have

$$\begin{aligned} \delta n &= -\int d^3\boldsymbol{x} \left( \frac{\delta n^{(1)}(\,\cdot\,|U)}{\delta \beta U(\boldsymbol{x})} \right)_{U=0} \delta U, & (4.4.52) \\ &= n \int d^3\boldsymbol{x} [n h(\,\cdot\, - \boldsymbol{x}) + \delta(\,\cdot\, - \boldsymbol{x})] \delta U = -n \left[ n \int d^3\boldsymbol{x}\, h(\boldsymbol{x}) + 1 \right] \delta U. & (4.4.53) \end{aligned}$$

---

[42] A thermodynamically more elementary derivation is as follows: Let us compute

$$k_BT \left.\frac{\partial n}{\partial P}\right|_V = \frac{k_BT}{V} \left.\frac{\partial N}{\partial P}\right|_V = \left.\frac{\partial N}{\partial \mu\beta}\right|_V \left.\frac{\partial \mu\beta}{\partial P}\right|_V.$$

We know $(\partial\mu\beta/\partial P)_{T,V} = \partial\mu\beta/\partial(k_BT \log \Xi/V)_{T,V} = \beta V/N$, so the fluctuation-response relation tells us

$$k_BT \left.\frac{\partial n}{\partial P}\right|_V = \frac{\beta}{N}\langle \delta N^2 \rangle$$

Now, we must compute $\langle N^2 \rangle$:

$$\langle N^2 \rangle = \int d\boldsymbol{r} d\boldsymbol{r}' \langle \rho^{(1)}(\boldsymbol{r}) \rho^{(1)}(\boldsymbol{r}') \rangle$$

Notice that $\rho^{(1)}(\boldsymbol{r})\rho^{(1)}(\boldsymbol{r}') = \rho^{(2)}(\boldsymbol{r}, \boldsymbol{r}') + \rho^{(1)}(\boldsymbol{r})\delta(\boldsymbol{r} - \boldsymbol{r}')$, we obtain

$$\langle N^2 \ \rangle = n^2 V \int g(\boldsymbol{r}) d\boldsymbol{r} + N = N \left( n \int h(\boldsymbol{r}) d\boldsymbol{r} + 1 \right) + N^2.$$

Thus, we arrive at (4.4.55).

From $P/k_BT = (1/V)\log\Xi$ and (4.4.26)

$$\frac{1}{k_BT}\delta P = \frac{1}{V}\int n d\boldsymbol{x}\delta U = n\delta U. \tag{4.4.54}$$

Taking the ratio of these formulas, we obtain the compressibility equation**q**

$$k_BT\left.\frac{\partial n}{\partial P}\right|_{V,T} = n\int h(\boldsymbol{r})d\boldsymbol{r} + 1. \tag{4.4.55}$$

The choice $A = \log\left[e^{\beta U(\boldsymbol{x})}n^{(1)}(\boldsymbol{x}|U)\right]$ and $B = n^{(1)}(\boldsymbol{x}|U)$ gives the *hypernetted chain* (HNC) closure

$$c(\boldsymbol{x}) = g(\boldsymbol{x}) - 1 - \log g(\boldsymbol{x}) - \beta\phi(\boldsymbol{x}). \tag{4.4.56}$$

With this closure the Ornstein-Zernike equation (4.4.4) may be solved for correlation functions. For short range interactions, the PY approximation seems better (more reliable).

> **Diagrammatics and closure approximations**
> We used a diagrammatic expansion to study imperfect gas. A similar approach is possible for liquid. The PY and HNC closures can be obtained diagrammatically by summing infinitely many diagrams up.[43] The HNC closure is obtained by summing all the diagrams that the current technology can sum up. The PY closure is obtained by summing up a smaller subset of diagrams. In any case, diagram approach has no physics. That is why this topic was not discussed at all.
>
> The PY closure was derived originally from a very different idea of producing collective coordinates for classical many-body systems. It is probably fair to say that the main project did not go anywhere, but a byproduct was the PY closure.[44]

What is an ideal liquid? We know ideal gas. It is deservingly called ideal gas, because real or imperfect gas may be obtained by perturbative approaches starting with (or expansion around) ideal gas. For gases, turning off all the interactions gives ideal gas. Even for liquids, this approach gives ideal gas, which is qualitatively different from liquid. One of the most important properties of liquid is coherence; it resists compression. The existence of the molecular repulsive core must be a vital element of liquid. Thus, it is a natural proposal to regard sufficiently dense hard core fluid as ideal liquid. Indeed, any liquid of spherical particles with short range attractive forces can be studied perturbatively around the hard core liquid

---

[43] G. Stell, "Percus-Yevick equation for radial distribution function of a fluid," Physica **29**, 517 (1963).

[44] J. K. Percus and G. J. Yevick, "Analysis of Classical Statistical Mechanics by Means of Collective Coordinates," Phys. Rev. **110**, 1 (1958). It is an interesting paper to read.

with an appropriate core size. The perturbation theory of liquid is quantitatively successful.[45,46]

> **Crowding effect**
> Dense fluid systems without any long range order is of vital importance in biology. Hard core fluids can be used to model colloidal solutions. The effect of excluded volume is quite serious in various processes in the cell, because 10%-40% of the volume is occupied by macromolecules. Crowding could enhance DNA folding and the function of chaperones.[47] It could protect bacteria and cells in general against large variation in tonicity of the environment and could even reverse the direction of reactions.[48]

## 4.5 Are there crystals?

As can be seen from (4.4.42) or (4.4.41) there is no anomalous behavior for the packing fraction $\eta$ up to unity. Since it has been demonstrated that the largest packing density in 3-space is the fcc or hcp densest close packing (for relevant comments see below), in reality there must be crystallization far below $\eta = 1$. Thus, it is clear that our liquid theory cannot describe the liquid-solid phase transition.

When a hard sphere fluid is compressed, crystalline order is found computationally, which is called the *Alder transition*.[49]

---

[45] D. Ben-Amotz and G. Stell, "Analytical implementation and critical tests of fluid thermodynamic perturbation theory," J. Chem. Phys. **119** 10777 (2003) may be a good starting point to look for references.

[46] ⟪**Bridge function**⟫ The integral equation theory is further developed by taking account of the so-called *bridge diagrams* (or bridge function) $B(r)$ defined as

$$h + 1 = \exp\{-\beta\phi + h - c + B\}.$$

$B = 0$ corresponds to the HNC approximation. How to write (approximately) $B$ in terms of $h, c$ and $\phi$ is the challenge of integral equation theory. J. M. Bomont and J. L. Bretonnet, "A self-consistent integral equation: Bridge function and thermodynamic properties for the Lennard-Jones fluid," J. Chem. Phys. **119**, 2188 (2003) may be a starting point to search references.

[47] protein folding catalysts

[48] R. J. Ellis, "Macromolecular crowding: obvious but underappreciated," Trends Biochem Sci **26**, 597 (2001) and a report of a meeting by G. Rivas et al., "Life in a crowded world — Workshop on the Biological Implications of Macromolecular Crowding," EMBO Rep **5**, 23 (2003) may be good starting references.

[49] ⟪**Does the Alder transition exist?**⟫ The original paper by Alder was about 2-dimensional hard disk system. It is known that there cannot be such a first-order phase transition in 2-space.

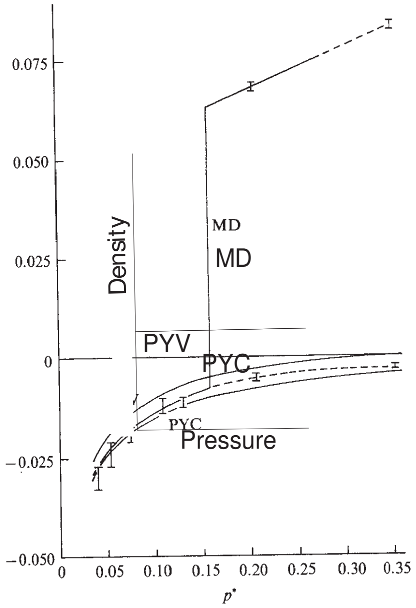


Fig. 4.5.1 PY equation of states and molecular dynamics simulation (based on B. J. Alder, J. Chem. Phys. **40**, 2724 (1964)).

Notice that this transition is solely entropy driven. The system is a hard core system, so the internal energy is only due to kinetic energy. Thus, crystal is solely driven by increasing entropy of configuration by regular packing. If we make a close random packing in a box and shake it, we would not hear rattling sound, but if we order the content (keeping the same number of particles), we would hear loud rattling noise.

Does the canonical or grand canonical partition function exhibit crystals? The fundamental question may be stated as follows:

> Suppose we have a 3D fluid system with short range interactions and sufficiently hard cores. If the temperature is sufficiently decreased and/or the density is sufficiently increased (in short, if the fugacity is sufficiently large), does the system lose translational symmetry?

In 2D we do not expect that a long-range order is possible due to its significant thermal fluctuation. We will discuss this point in detail in the next chapter. However, at $T = 0$ for the classical case, there is a proof that for the ordinary fluid systems the lowest energy state is a triangle lattice.[50] In 3D it is still an outstanding open problem of mechanics (and statistical mechanics) to show that the lowest

---

Therefore, the transition found by computer simulations was a finite-size effect.

[50] F. Theil, "A Proof of Crystallization in Two Dimensions," Commun. Math. Phys., **262**, 209-236 (2006).

energy state is periodic.

A related question is the packing and tiling problems.[51] One might easily guess that the crystallization problem is intimately related to the packing problem of hard spheres. It has been proved that the fcc or hcp packing density $\pi/\sqrt{18} \sim 0.74048\cdots$ is the highest density possible in 3-space.[52] However, the 'ground state' is uncountably degenerate, because aperiodic stacking of 2D hexagonal lattice is possible while maintaining the maximum possible density. It is believed that adding weak attractive forces will lift the degeneracy and the periodic structure becomes the energy minimum state. Although we have seen that hard core interactions alone cannot make any periodic structure, still they have almost decisive effects on crystal structures as Kihara demonstrated with his magnet models.[53]

---

[51] A readable review article is: C. Radin, "Low temperature and the origin of crystalline symmetry," Internat. J. Mod. Phys. B **1**, 1157 (1987). For a very recent related article, see J. H. Conway and S. Torquato, "Packing, tiling, and covering with tetrahedra," Proc. Natl. Acad. Sci., **103**, 10612 (2006).

[52] ⟪**The Kepler problem**⟫ This was first conjectured by Kepler. T. C. Hales, "A proof of the Kepler conjecture," Ann. Math., **162**, 1065-1185 (2005). For ellipsoids, this density can be larger: $(24\sqrt{2} - 6\sqrt{3} - 2\pi)\pi/72 \simeq 0.7533355\cdots$ (A. Bezdek and W. Kuperberg, in *Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift* (ed. P Gritzmann and B Sturmfels) (AMS Providence, RI 1991) p71).

[53] T. Kihara, *Intermolecular forces* (Wiley, 1978).

## Exercises for Chapter 4 (rto Maxwell, rto Tonks, rto Debye-Hückel)

**4.1**. [Kac potential]
There is an imperfect classical gas with a binary potential $\phi$ given by

$$\phi(|\boldsymbol{r}|)=\begin{cases}\infty & \text{if } |\boldsymbol{r}|<a,\\ -\varepsilon/l^3 & \text{if } a\le|\boldsymbol{r}|<l,\\ 0 & \text{otherwise.}\end{cases} \tag{4.P.1}$$

Here, $\varepsilon>0$, $a$ is a length scale of atom size, and the $l\to\infty$ limit is taken. (This is an example of the Kac potential.)
(1) Compute the second virial coefficient (in the $l\to\infty$ limit).
(2) Compute the Joule-Thomson coefficient $(\partial T/\partial P)_H$, where $H$ is enthalpy. The reader may assume that the heat capacity $C_P$ under constant pressure is a constant and is known,

**4.2** [van der Waals equation of state]r
(1) Show that the critical point is defined by

$$\left.\frac{\partial P}{\partial V}\right|_T=\left.\frac{\partial^2 P}{\partial V^2}\right|_T=0. \tag{4.P.2}$$

(2)r For the van der Waals equation of state, find the universal ratio $P_cV_c/k_BT_c$.
(3) Obtain the reduced equation of state $P_r=f(V_r,T_r)$ for the van der Waals gas. Here, $P_r=P/P_c$, $V_r=V/V_c$ and $T_r=T/T_c$ are reduced variables. [The reader can work with a 1 mole gas.]
(4) Near the critical point $P_r-1$ may be expanded in powers of $T_r-1$ and $n_r-1$, where $n_r=1/V_r$ is the reduced number density. Find the coefficients $A$ - $C$ (we will see a close relation of this to the Landau theory of phase transition later).

$$P_r-1=A(T_r-1)+B(T_r-1)(n_r-1)+C(n_r-1)^3+\cdots. \tag{4.P.3}$$

(5) For hydrogen gas $H_2$, $b=26.61$ cm$^3$/mol. This volume can be interpreted as the smallest volume that the molecules can be squeezed into. Assuming that Avogadro's constant is known, estimate Planck's constant (use dimensional analysis to guess the atom size in terms of electron charge $e$, mass $m$, $h$ and $4\pi\varepsilon_0$).

**4.3** [The free energy of the van der Waals gas]
The Helmholtz free energy of the van der Waals gas may be expressed as

$$A=-Nk_BT\left\{\log\left[\frac{n_Q}{N}(V-Nb)\right]+1\right\}-\frac{aN^2}{V}. \tag{4.P.4}$$

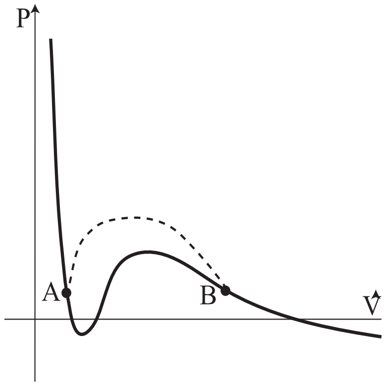
(1) Comparing this with the free energy formula for the ideal gas, explain why this form is natural.

(2) Compute the internal energy and the entropy of the van der Waals gas.

**4.4** [Thermodynamically respectable derivation of Maxwell's rule]
If the temperature is sufficiently low, the $PV$-curve given by the van der Waals equation of state implies

$$\frac{\partial P}{\partial V} = -\frac{Nk_BT}{(V - N\sigma)^2} + a\frac{N^2}{V^3} > 0. \tag{4.P.5}$$

That is, it is thermodynamically unrealizable. Actually, gas-liquid coexistence occurs when this 'unphysical behavior' happens, and the coexistence temperature $T$ is determined by the Maxwell rule. This is what Maxwell proposed and an 'explanation' was given in the text but was with a remark that the argument is an abuse of thermodynamics. Many textbooks argue that Maxwell's rule cannot be derived thermodynamically properly, because the argument in the text (the usual one) utilizes thermodynamics where the states are unstable. However, it is possible to avoid this abuse and still we can thermodynamically demonstrate Maxwell's rule. The coexistence condition for phase A and phase B is the agreement of $P$, $T$ and $\mu$. $\mu_B(T, P) - \mu_A(T, P)$ of the difference of the Gibbs free energy must be computable along the path in the phase diagram through only stable phases (that is the broken curve in the following figure).



Since

$$G = E - ST + PV, \tag{4.P.6}$$

if we compute $E_B - E_A$ and $S_B - S_A$, then $G_A = G_B$ allows us to compute the difference of $PV$, that is, $P(V_A - V_B)$.
(1) Compute $E_B - E_A$.
(2) Compute $S_B - S_A$.
(3) Since $G_B - G_A = 0$, these results allow us to compute $P(V_B - V_A)$. Confirm that

this and the result obtained by the naive abuse of thermodynamics:

$$\int_A^B PdV \tag{4.P.7}$$

agree.

**4.5** [Grand canonical approach to 1D van der Waals gas]
Let us study the 1D Kac model with the aid of the grand canonical approach.
(1) If there are $N$ particles in the container of volume $V$, the canonical partition function reads

$$Z_N(V) = \int_{(N-1)\sigma}^{V-\sigma} dx_N \cdots \int_{\sigma}^{x_3-\sigma} dx_2 \int_0^{x_2-\sigma} dx_1 \int dp_1 \cdots dp_n e^{-\sum_{i=1}^N p_i^2/2mk_BT + aN^2/k_BTV}. \tag{4.P.8}$$

After checking the formula is correct, actually compute this.
(2) Using the result of (1) write down the grand partition function (you cannot perform the summation in a closed form).
(3) The grand partition function written down in (2) has the following structure:

$$\Xi = \sum_{N=0}^{M} e^{VA(N/V)}, \tag{4.P.9}$$

where $M$ is the maximum number of particles we can push into the volume $V$. Show that if the temperature is sufficiently high, there is only one $n = N/V$ that maximizes $A(n)$. Also demonstrate that if the temperature is sufficiently low, there can be three extrema for $A(n)$.
(4) What do you expect the grand partition function looks like, if $n$ that maximizes $A(n)$ are not unique?
(5) There is a text book which writes explicitly as follows:

$$\Xi = e^{\beta PV} + e^{\beta P'V}. \tag{4.P.10}$$

Here, we have assumed that $A(n)$ have two maxima, and the two terms correspond respectively to the two maxima. Is this correct?

**4.6** [Hard sphere fluid]
The virial equation of state for a fluid interacting with 2-body potential reads

$$\frac{P}{nk_BT} = 1 - \frac{2\pi}{3}\beta n \int_0^\infty \phi'(r)g(r)r^3dr. \tag{4.P.11}$$

Using this formula, derive the equation of state for the hard sphere fluid:

$$\frac{P_H}{nk_BT} = 1 + \frac{2\pi}{3} n g(\sigma)\sigma^3, \tag{4.P.12}$$

where $\sigma$ is the diameter of the sphere. Strictly speaking, $g(\sigma)$ is $\lim_{r\searrow\sigma} g(r)$.

**4.7** [Internal energy of fluid interacting with binary forces]
(1) Suppose the interactions among spherical particles can be expressed in terms of the two-body interaction potential $\phi(\boldsymbol{r}) = \phi(r)$. Write down its internal energy in terms of $E$ $\phi$, the number density $n$ and the radial distribution function $g(r)$.
(2) Obtain the internal energy of the Kac fluid (i.e., the fluid interacting via the Kac potential). Set $\sigma = 1$.
(3) Using the virial equation of state

$$P/nk_BT = 1 - \frac{2\pi}{3} n\beta \int_0^\infty \phi'(r)g(r)r^3 dr, \tag{4.P.13}$$

obtain the augmented van der Waals equation of state for a Kac fluid:

$$\beta P = \beta P_H + \frac{n^2}{2}\beta \int d^3\boldsymbol{r}\, \phi(\boldsymbol{r}). \tag{4.P.14}$$

Notice that the radial distribution function of the Kac fluid is the same as the had sphere fluid with the same number density.[54]

**4.8** [Functional differentiation]
Compute the following functional derivative.

$$\frac{\delta}{\delta\psi(\boldsymbol{x})} \int d^3\boldsymbol{r} \left[ \frac{1}{2}(\nabla\psi(\boldsymbol{r}))^2 - \frac{1}{2}\tau\psi^2(\boldsymbol{r}) + \frac{1}{4}\psi^4(\boldsymbol{r}) \right], \tag{4.P.15}$$

where $\tau$ is a constant.
(2) [Green's function and functional differentiation] Consider a differential equation (partial or ordinary) $L\varphi = f$, where $L$ is a linear differential operator acting on the functions of $x$.[55] Show that $\delta\varphi/\delta f$ is the Green's function for the initial boundary value problem defined by the linear operator $L$. (This is a problem immediately solved by inspection, but the fact is not meaningless. The method of Green's function is actually a method to solve a differential equation by the first order functional Taylor expansion approach explained in the text; in this case the problem is linear,

---

[54] However, we cannot derive the Maxwell's rule requirement.

[55] If $(Lf)(x)$ is determined by the value of $f$ at $x$ and the values of various derivatives of $f$ at $x$, $L$ is called a differential operator.

so the method gives an exact solution.)
(3) Regard the entropy $S$ of a fluid interacting with the binary potential $\phi$ as its functional, and express the functional derivative

$$\frac{\delta S}{\delta \phi(\boldsymbol{r})} \tag{4.P.16}$$

in terms of the radial distribution function (and its appropriate partial derivatives). It may be easy to compute the corresponding functional derivative of the Helmholtz free energy.

**4.9** [Functional derivatives of canonical partition function]
Redo the calculations (4.4.25)-(4.4.27) using the canonical formalism; this is slightly easier than the grand canonical approach given in the text.

**4.10** [Debye-Hückel theory]
Let us perform the functional Taylor approximation approach explained in Section 4.4 with $A = n$ and $B = U$.
(1) Within this approach find the equation governing the radial distribution function $g$.
(2) Obtain the Fourier transform of the in direct correlation function in the present approximation.
(3) Let $\phi$ be the Coulomb potential. Its Fourier transform may be written as $\phi_k = Q/k^2$. What is the functional form of the indirect correlation function $h(r)$?

**4.11** [Toy integral equation for liquid]
Let us make a prototype closure for the radial distribution.
(1) Make the first order approximation (i.e., the first order functional Taylor expansion approximation) of $n^{(1)}(\boldsymbol{x}|U)$ in terms of the Boltzmann factor $\exp[-\beta U(\boldsymbol{x})]$.
(2) What is the direct correlation function?
(3) Find the Fourier transform of the indirect correlation function.
(4) Find the equation of state with our approximation with the aid of compressibility or its reciprocal. Assume that the diameter of the hard core is $a$.

**4.12**. [*Scaled particle theory* of hard core fluid][56]
As we know well by now, for a spherical hard core fluid

$$P/nk_BT = 1 + \frac{2\pi}{3} n\sigma^3 g(\sigma), \tag{4.P.17}$$

---

[56] R. J. Baxter, in *Physical Chemistry an advanced treatise* volume VIIIA Liquid State (edited by H. Eyring, D. Henderson and W. Jost, Academic Press 1971) Chapter 4, Section VIII.

where $\sigma$ is the diameter of the spherical core. Therefore, to know the hard core equation of state we need $g$ only at $r = \sigma$.

Let $p_0(r)$ be the probability of observing a bubble of radius $r$. Let $nG(r)$ be the expected number of the centers just outside the bubble ($nGdr$ is the expected number of particle centers in the spherical shell between $r$ and $r + dr$). When the bubble is of radius $\sigma$, it just behaves as the exclusion zone by the hard sphere at the origin. Therefore,

$$g(\sigma) = G(\sigma). \tag{4.P.18}$$

We have only to determine $G$ to know the hard-core fluid pressure.
(1) Derive

$$p_0(r + dr) = p_0(r)[1 - 4\pi n r^2 G(r) dr]. \tag{4.P.19}$$

That is,

$$\frac{d}{dr} \log p_0(r) = -4\pi n r^2 G(r). \tag{4.P.20}$$

(2) We can determine $G$ for very large $r$. According to the fluctuation theory, the probability of fluctuation that creates a bubble of radius $r$ may be written in terms of the reversible work $W(r)$ required to make it. Therefore,

$$p_0(r) = e^{-\beta W(r)}. \tag{4.P.21}$$

Using this and the thermodynamic result for large $r$ (i.e., for the usual macroscopic bubble!)

$$dW(r) = PdV + f[1 - (2\delta/r)]dA, \tag{4.P.22}$$

where $A$ is the surface area of the bubble, and $f(1 - 2\delta/r)$ is the surface tension of the curved surface of mean curvature $1/r$. Using (6.4.119)-(6.4.121), find $G(r)$ as a function of $r$.
(3) If $r < \sigma/2$, only 1 particle center can come in the bubble. What is this probability? This must be $1 - p_0(r)$ for $r < \sigma/2$.
(4) Determine $G(r)$ for $r < \sigma/2$.
(5) Unfortunately, $G(r)$ is not a smooth function, but it is known that it is continuously differentiable at $r = \sigma/2$. Let us make an approximation that the unknown parameters $f$ and $\delta$ may be determined by matching $G$ obtained in (2) and in (4) at $r = \sigma/2$ smoothly (match $G$ and $G'$ there). Derive, under this approximation,

$$\frac{P}{nk_BT} = \frac{1 + \eta + \eta^3}{(1 - \eta)^3}, \tag{4.P.23}$$

where $\eta$ is the packing density: $\eta = \pi\sigma^3 n/6$ as usual.
[This is identical to PY-C! Furthermore, $f$ obtained is quite reasonable.]

**4.13** [Quantum effect on the second virial coefficient][57]
The second virial coefficient for a spherical symmetrical particle is, classically,

$$B = 2\pi \int_0^\infty (1 - e^{-\beta\phi(r)}) r^2 dr. \tag{4.P.24}$$

Its quantum version should be obtained by replacing the Boltzmann factor $e^{-\beta\phi}$ with the diagonal element of the 2-body density operator $\rho(r)$:

$$B = 2\pi \int_0^\infty (1 - \mathcal{N}\rho(r)) r^2 dr, \tag{4.P.25}$$

where $\mathcal{N}$ is the normalization constant to make $\mathcal{N}\rho(r) \to 1$ in the $r \to \infty$ limit.[58]

Thus, the calculation has been reduced to that of $\rho$. Let us perform this through small steps. The difference of fermions and bosons can be ignored except for helium below 25 K. Therefore, we totally ignore the effect of particle symmetry on the wave function. The following calculation was done for the first time by E. Wigner in the 1930s. For $^4$He, below 50 K the quantum correction increases the classical value by about 50%, but 10 % around 100 K, so the quantum correction is not very small.
(1) Let us assume that two-body Hamiltonian to be

$$H = -\frac{\hbar^2}{2m}(\Delta_1 + \Delta_2) + \phi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|), \tag{4.P.26}$$

where $\Delta_i$ is the Laplacian wrt the position coordinates of the $i$-particle. The two-body density operator is given by $\rho = e^{-\beta H}$. Show that the matrix element $\langle \boldsymbol{r}_1, \boldsymbol{r}_2 | \rho | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle$ of the 2-body density operator satisfies the following equation:

$$\frac{\partial}{\partial\beta}\langle \boldsymbol{r}_1, \boldsymbol{r}_2 | \rho | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle = -H \langle \boldsymbol{r}_1, \boldsymbol{r}_2 | \rho | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle. \tag{4.P.27}$$

If we use the normalized eigenket $|i\rangle$ ($H|i\rangle = E_i|i\rangle$) of $H$, we can write

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2 | \rho | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle = \sum_i \langle \boldsymbol{r}_1, \boldsymbol{r}_2 | i \rangle e^{-\beta E_i} \langle i | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle. \tag{4.P.28}$$

---

[57] This is based on T. Kihara, *Molecular forces.*

[58] As we will see below, the diagonal element of the position representation of the density operator is proportional to the probability of finding particles, so it is proportional to the Boltzmann factor semiclassically. Therefore, normalizing the density operator appropriately, we can interpret it as a quantum statistical extension of the spatial Boltzmann factor.

Here, the summation may include integration.
(2) For the case with $\phi = 0$ (i.e., for the ideal gas case) obtain $\langle \boldsymbol{r}_1, \boldsymbol{r}_2 | \rho | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle$. As already noted, you can totally forget about the particle exchange symmetry. You must specify the initial and the boundary conditions correctly to solve the above parabolic equation.
(3) Let us introduce the deviation $\Psi$ from the case without interactions as

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2 | \rho | \boldsymbol{r}_1', \boldsymbol{r}_2' \rangle = \left( \frac{2\pi\hbar^2\beta}{m} \right)^{-3} \exp\left[ -\frac{m}{2\hbar^2\beta}[(\boldsymbol{r}_1 - \boldsymbol{r}_1')^2 + [(\boldsymbol{r}_2 - \boldsymbol{r}_2')^2] + \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2; \boldsymbol{r}_1', \boldsymbol{r}_2') \right]. \tag{4.P.29}$$

Note that $-k_BT\Psi$ corresponds to the 'quantum-corrected two-body interaction. $\Psi$ satisfies the following equation:

$$\phi + \frac{\partial \Psi}{\partial \beta} + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}_1') \cdot \nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}_2') \cdot \nabla_2]\Psi = \frac{\hbar^2}{2m}[(\nabla_1\Psi)^2 + \Delta_1\Psi + (\nabla_2\Psi)^2 + \Delta_2\Psi]. \tag{4.P.30}$$

This equation can be obtained from (6.4.143) by substituting (6.4.145); nothing special has not been done at all.

To obtain the quantum correction we expand $\Psi$ as

$$\Psi = \Psi_0 + \frac{\hbar^2}{2m}\Psi_1 + \cdots \tag{4.P.31}$$

and then introduce this into (6.4.146). Requiring the order by order agreement of the substituted result, we get

$$\phi + \frac{\partial}{\partial \beta}\Psi_0 + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}_1') \cdot \nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}_2') \cdot \nabla_2]\Psi_0 = 0, \tag{4.P.32}$$

$$\frac{\partial}{\partial \beta}\Psi_1 + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}_1') \cdot \nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}_2') \cdot \nabla_2]\Psi_1 = (\nabla_1\Psi_0)^2 + \Delta_1\Psi_0 + (\nabla_2\Psi_0)^2 + \Delta_2\Psi_0, \tag{4.P.33}$$

etc. First, we must solve the zeroth order equation. Show that the diagonal element of the only meaningful solution is $\Psi_0(\beta, \boldsymbol{r}_1, \boldsymbol{r}_2; \boldsymbol{r}_1, \boldsymbol{r}_2) = -\beta\phi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$.
(4) We have only to obtain the diagonal element of $\Psi_1$. Solve the simplified equation that can be obtained by taking the diagonal limit of (6.4.149):

$$\frac{\partial}{\partial \beta}\Psi_1 = \lim_{\{\boldsymbol{r}_i'\} \to \{\boldsymbol{r}_i\}} \left[ (\nabla_1\Psi_0)^2 + \Delta_1\Psi_0 + (\nabla_2\Psi_0)^2 + \Delta_2\Psi_0 \right]. \tag{4.P.34}$$

As can immediately be seen, to compute the RHS of this equation, we cannot use $\Psi_0(\beta, \boldsymbol{r}_1, \boldsymbol{r}_2; \boldsymbol{r}_1, \boldsymbol{r}_2) = -\beta\phi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$ that is already in the diagonal limit. The derivative must be computed from the original expression of $\Psi_0$. Find $\lim_{\{\boldsymbol{r}_i'\} \to \{\boldsymbol{r}_i\}} \nabla_1\Psi_0$

and $\lim_{\{\boldsymbol{r}'_i\}\to\{\boldsymbol{r}_i\}} \nabla_2\Psi_0$.
(5) To botain $\Delta_1\Psi_0$ we could perform a similar calculation. However, our purpose is not to practice calculation (although this practice is also meaningful), the needed result is provided:

$$\Delta_i\Psi_0 = -\frac{1}{3}\beta\Delta_i\phi, \tag{4.P.35}$$

where $i = 1, 2$.

Obtain the diagonal element $\langle\boldsymbol{r}_1, \boldsymbol{r}_2|\Psi_1|\boldsymbol{r}_1, \boldsymbol{r}_2\rangle$. Then compute the quantum correction to the second virial coefficient to order $\hbar^2$. Show that the correction is positive for any $T$ (as stated for helium, quantum correction always increases the second virial coefficient).
(6) The quantum correction starts with the term of $O[\hbar^2]$. Why, or why is't there any correction of odd order in $h$?

# Chapter 5

# Phase Transitions

For systems with interactions various 'phases'are often found other than the disordered phase that is usually found at high temperatures. The state of a system may depend on temperature, pressure and other thermodynamic variables. A 'phase diagram' is a 'territory map' of various phases in the space spanned by these thermodynamic variables. Our 'common-sense phases' such as liquid phase or solid phase usually occupy a large domain in this diagram. The boundary of the territory may or may not be well-defined, but to have a rough understanding of this map should be the first step to understand the character of the system. To this end, we should first understand where the territory boundaries run and how these boundaries are characterized. If distinct phases may be definable at all, they must have distinct thermodynamic characteristics. Therefore, at the territory boundary (where a phase changes to another phase) the change of thermodynamic variables must be differently from that inside the territory. Phase boundaries and phases affected by the existence of nearby phase transitions occupy only a very small portion of the whole phase diagram, but the importance of understanding phase transitions should be clear.

After an overview section, two sections are devoted to explain thermodynamic limit to characterize phase transitions precisely. These two sections may be skipped for the first reading without losing an overall picture. For a phase other than the high temperature disordered phase to exist, the phase must be stable against thermal fluctuation. If the spatial dimension is not large (not larger than four) the effect of fluctuation is decisive. The fluctuations can be spatially strongly correlated fluctuations near the phase boundaries, so we need a generalization of the central limit theorem (renormalization) to handle them. However, if we wish to discuss a phase sufficiently away from the phase boundaries, the correlation of fluctuations become relatively insignificant. Therefore, we may understand the qualitative features of phases without paying due attention to fluctuations, or we may average fluctuations out (mean field theories). Therefore, in this Chapter after surveying the effects of fluctuations on phases and phase transitions, the idea of renormalization group theory is introduced intuitively, and then mean field theory is outlined. Roughly speaking, in order to understand the phase diagram, we have only to understand where renormalization group theory is needed and where mean field theory is effective.

## 5.1 Does a phase order?

We have already seen two exactly solvable models exhibiting phase transitions (the ideal Bose gas and the 1D Kac potential system). Although not without approximation we have also studied fluid systems which are more realistic than the above examples. Let us introduce a lattice model of ferromagnets that is not so artificial but perhaps more tractable than fluids. It is the *Ising model* introduced by W. Lenz.[1] The model consists of spins $\{s_i\}$, where each spin $s_i \in \{-1, +1\}$ with $i$ indicating lattice points. We consider here simple regular lattices with a finite coordination number such as the cubic lattice. The system Hamiltonian (under an external magnetic field; the intrinsic Hamiltonian + interaction potential with the magnetic field, precisely speaking) reads

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_i s_i, \tag{5.1.1}$$

where the summation is over all the nearest neighbor spin pairs on the lattice ($\langle i, j\rangle$ implies an adjacent spin pair of spin $i$ and $j$), $J > 0$ is called the *coupling constant*, and $h$ is the magnetic field.

---

[1] **⟨⟨W. Lenz⟩⟩** W. Lenz (1888-1957) is different from the Lenz of Lenz's law. He was an important figure in the early development of quantum mechanics, known for the Lenz vector (Laplace-Runge-Lenz vector relevant to the $O_4$ symmetry of the Kepler problem). He was a student of and, later, a long-time assistant to Sommerfeld. Pauli, Jordan and others were his assistants. He was an important figure in the development of theoretical physics in Germany.

**Spin-spin interaction**
Ferromagnets result from electron exchange interactions aligning atomic magnets. Atomic magnets may be understood as magnetic moments associated with electron angular momenta.[2] This purely quantum mechanical exchange interaction proposed by Heisenberg gives an interaction energy proportional to

$$-J\boldsymbol{s}\cdot\boldsymbol{s}' \tag{5.1.2}$$

between adjacent atomic magnetic moments $\boldsymbol{s}$ and $\boldsymbol{s}'$, where $J$ ($> 0$) is a constant called the *coupling constant*.

If the crystal has a direction in which the magnetic moments tend to point (called the *easy direction* or *easy axis*), then we may approximate the magnetic moment as a scalar and (5.1.2) becomes

$$-Jss'. \tag{5.1.3}$$

If spins $s$ and $s'$ take only two values, say $\pm 1$, they are called *Ising spins*.[3]

If the temperature is high, due to large thermal fluctuation, it is not easy for neighboring spins to be parallel with each other. Therefore, at sufficiently high temperatures spin directions are various, so a disordered phase should be the equilibrium state in which the expectation value of the spin $\langle s_i \rangle$ is zero (*disordered state*). To make a phase in which many spins align in one direction, say, up (ferromagnetic phase) where $\langle s_i \rangle$ is nonzero, the aligned phase must be stable against thermal noise. If such a state is realized at sufficiently low temperatures, there must be a phase transition from a disordered phase to an ordered phase at some temperature $T_c > 0$, where $\langle s_i \rangle$ changes from zero to some nonzero value.

There is a way to understand why $T_c > 0$ in spaces with dimension larger than or equal to 2 for the Ising model.**r**

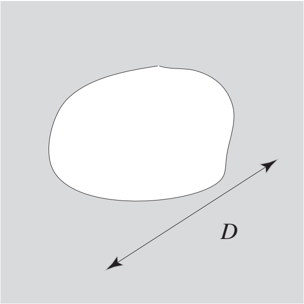


Fig. 5.1.1 Peierls' argument illustrated for 2-space. The gray zone is the up spin ocean, and the white zone is the down spin island to be created by thermal fluctuations. The needed energy is proportional to the size $D$ (proportional to the length of the circumference).

Let us consider a 2-Ising model (two-dimensional Ising model), and consider the

---

[2] In a crystal due to the crystal field orbital angular momenta disappear in the ground state of the atom, so the spin angular momenta contribute mainly, quite different from free ions.

[3] Strictly speaking, this is the case with $S = 1/2$.

state with all spins down. Let us take a circular island of diameter $D$, and flip all the spins on it. If this spin flipping costs no macroscopic energy, then thermal fluctuation could afford such flipping, and we cannot expect any stable ordered phase at finite temperatures. The energy required for this flipping is $\pi D \times 2J$. Therefore, if $D$ is large, this is impossible at sufficiently low temperatures;$_{\mathbf{q}}$ we may expect ordering of spins at sufficiently low temperatures. Thus, we may conclude that 2-Ising model can exhibit order-disorder phase transition at $T_c > 0$.[4] This argument, called *Peierls' argument*,$_{\mathbf{r}}$ can actually be mathematically justified. Stated in a precise form, it says that even if the all up-spin boundary recedes to infinity, the probability of the spin at the center of the system to be up remains strictly larger than 1/2 (see the next section and an exercise at the end of this Chapter). It is clear that Peierls' argument can be applied to $d$-space with $d \geq 2$. That is, $d$-Ising model ($d \geq 2$) exhibits an order-disorder phase transition at a finite temperature, $T_c > 0$.

Let us repeat the same argument as above in 1-space.$_{\mathbf{r}}$ The energy cost to flip all the spins on the segment of length $D$ upward in the ocean of down spins is always $4J$ irrespective of $D$, because the boundaries are just two points. This is a microscopic energy; only 'finite patience' is needed to obtain it even if the temperature is very low (but positive). Therefore, there should not be any phase transition at a finite temperature. To be precise, we must take account of the entropy$_{\mathbf{q}}$ due to the location of the island, which is proportional to the logarithm of the system size, and is definitely larger than $\log D$. Therefore, the free energy cost $\Delta A$ to create the island must be smaller than $4J - k_BT \log D$. This implies that as long as $T > 0$, the boundary is formed spontaneously.

We have realized that in lower dimensional spaces fluctuation effect is serious. 1-Ising model cannot order at any positive $T$. As we will discuss later, 2-Ising model can order below $T_c$, but up-spin phase and down-spin phase cannot coexist (we say there are only two pure Gibbs states as we will learn later), because the interface

---

[4] ⟪**Effect of boundary entropy**⟫ In reality, the phase boundary is microscopically rugged, so the boundary is stabilized due to the entropic effect. In other words, the free energy required to form the boundary is reduced and the boundary becomes easier to form. However, this 'surface entropy' must be proportional to the surface area of the (coarse-grained) boundary (the extensivity in the space one dimension lower than the bulk space), so it is proportional to $D$. Let us write the proportionality constant to be $c$. The free energy cost to make an island in Fig. 5.1.1 becomes $\Delta A = 2\pi JD - cTD$. If the temperature is sufficiently low, this certainly becomes positive, so the rough argument given in the text is justified.

There is also the contribution of the positional freedom of the island (i.e., where to position the center of the island) to the entropy, but it is proportional to the logarithm of the system volume, so compared to $D$, which is proportional to $V^{1-1/d}$ in the general case, we may ignore it as long as $d > 1$. We cannot ignore this in 1-space as noted below.

fluctuates violently. Nontrivial critical behavior mentioned in 2 and 3D for the Ising model is due to fluctuation. The fluctuation effect on phase transition is crucial.

Why spatial dimensionality is important can be understood if we consider how vulnerable the 'communication paths' between two lattice points are against thermal fluctuations. Spatially separated two Ising spins can unite to resist thermal fluctuations only when they can communicate well (that is, one spin must clearly know whether the other spin is up or down). If there are many communication routes, then fluctuations cannot easily disrupt the communication between the spins, so they can work together. In particular, in 1-space, any two distinct spins have only one communication route. If this is destroyed by a fluctuation, that is the end of the collaboration. Therefore, the long range order in 1-space is vulnerable to fluctuations. In 2-space, there are sufficiently many routes for two spatially separated Ising spins to communicate, so fluctuations cannot disrupt the cooperation of spins that are even far apart, if the temperature is low enough.

Another way to understand the spatial dimensionality is to ask how the effect of disturbance at one point can propagate in space. This is to observe the response of a phase against spatial perturbation. Since response and fluctuation are closely related, this is also a natural approach to understanding the nature of the phases. If the effect of a localized disturbance is global, then the stability of the phase may be questioned.

To study the global response of the system to a localized disturbance, we have only to look at the system from distance. Even in the case of Ising ferromagnets in an ordered state not all the spins point the same direction, so if we look at the magnet from distance, it looks 'gray.' That is, it is sensible to describe the order in terms of a continuous field $\varphi$ (a course-grained order parameter field; for example, for the all up phase $\varphi = +1$, for the all down phase $\varphi = -1$, and for other phases the values are appropriately assigned to interpolate these values). We expect that a spatially uniform equilibrium state is stable, so we can expect that there is a restoring force against disturbances that could destroy the spatial uniformity $\varphi =$ const. The restoring force must be a linear force for small disturbances, so it must be proportional to $-\Delta\varphi$ ($\Delta$ is the Laplacian; see the fine letters below). Then, the fundamental solution[5] of the Laplace equation $\Delta\varphi = 0$ is the key to understanding the disturbance effect. It behaves roughly as $1/r^{d-2}$ in $d$-space. An ordered phase is clearly very vulnerable in 1-space, and stable in 3-space. If we pay attention to the interface, it is one dimensional in 2-space, so it is unstable. This explains why the interface in 2-Ising model fluctuates violently as we see in Sect. 5.4.

---

[5] The fundamental solution satisfying the given auxiliary conditions is the Green's function.

**Intuitive understanding of Laplacian**r
If we discretize the 1-Laplacian (i.e., $d^2\varphi/dx^2$) using the simple Euler scheme, it reads

$$-\frac{d^2\varphi}{dx^2} \simeq \frac{2}{\delta x^2}\left(\varphi(x) - \frac{\varphi(x+\delta x)+\varphi(x-\delta x)}{2}\right). \tag{5.1.4}$$

That is, if we denote the nearest neighbor average of $\varphi$ at $x$ as $\langle\langle\varphi(x)\rangle\rangle$

$$-\Delta\varphi(x) \propto \varphi(x) - \langle\langle\varphi(x)\rangle\rangle. \tag{5.1.5}$$

This is true in any dimensional space.

$\varphi$ is a *harmonic function* if and only if $\Delta\varphi = 0$. Our interpretation immediately tells us that harmonic functions cannot have any local extremum. Furthermore, the famous *spherical average theorem* (for a harmonic function its average value on a sphere is equal to the value at the center of the sphere) is intuitively clear.

In the above Peierls' argument we assumed that the system was sufficiently large. If not, needless to say, $D$ cannot be macroscopic in Fig. 5.1.1, so no stable ordered phase can exist. As we have already learned, to formulate thermodynamics precisely, we must take the thermodynamic limit. Therefore, we should also take the same limit to discuss phase transitions without ambiguity. Let us outline how to take the thermodynamic limit and its consequences in the next two sections.

## 5.2 Gibbs measure and phase transition

It is almost meaningless to discuss solid and liquid phases if the system is a cluster of 10 molecules. Needless to say, to state clearly that the system is in a solid state, it must be sufficiently large. How large is large? To avoid such an awkward question, the simplest way is to take the large system size limit. The phase transition becomes sharp in larger systems, so thermodynamic limit, formulating the large system-size limit precisely, has special significance in the present context: it gives a possibility to characterize phase transition as a mathematical singularity or a bifurcation point.

How can we study the thermodynamic limit of a system?r The Hamiltonian in this limit is meaningless; it is generally a non-convergent infinite sum. For example, see the Ising Hamiltonian; it is an infinite sum of ±1. Consequently, the Boltzmann factor is almost always 0 or ∞.

Because the purpose of thermodynamic limit is to study the bulk properties of a system, we should take the van Hove limit.[6] Thus, in order to consider thermodynamic limit we prepare nested sequences of volumes whose surfaces are not very complicated. A finite system must have, usually, a boundary, so boundary conditions must be imposed. The reader might think the boundary conditions lose significance in the thermodynamic limit, but this is not always the case. We already know that for 2-Ising model, the up-spin boundary condition far away indeed affects what we observe around us, if $T < T_c$. In most interesting cases boundary effects are crucial.

In this section a procedure to take the thermodynamic limit is outlined with the aid of a lattice model with finite-range interactions. Suppose we wish to study an infinite lattice $L$. We prepare nested sequence $\{V_i\}$ of (van Hove) volumes converging to $L$ (i.e., $V_1 \subset V_2 \subset \cdots \subset L$).

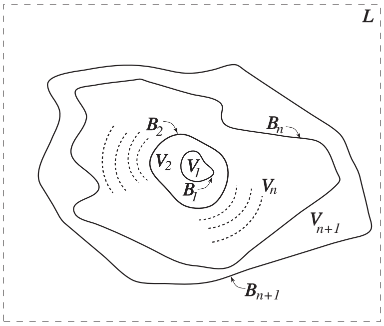


Fig. 5.2.1 An increasing nested sequence of finite systems to define thermodynamic limit. $L$ is the infinite lattice and $V_i$ are finite systems with van Hove volumes with explicit boundary conditions $B_i$ (= the minimal sufficient condition to fix the system energy uniquely when the state in $V_i$ is given).

For each finite volume system a boundary condition $B$ must be specified.[7] Let us write the volume-boundary condition pair as $(V, B)$. Let us make the canonical distribution for $(V, B)$:

$$\mu_{V,B} = \frac{1}{Z_{V,B}} e^{-\beta H_V(B)}, \tag{5.2.1}$$

where $H_V(B)$ is the Hamiltonian for the system with the volume $V$[8] and the boundary condition $B$ (thus $H_V(B)$ includes the surface interaction terms). Make the totality of $\mu_{V,B}$ choosing all possible $V \subset L$ and $B$. Then, we take a sequence $\{\mu_{V_i,B_i}\}$ that converges to a probability measure $\mu$ (defined on the set of microstates on $L$). The

---

[6] Roughly speaking, this implies that we take the large system size limit so that the surface area $\times L$ scales as the system volume, where $L$ is the representative length scale of the system.

[7] For the Ising model the interaction is nearest neighbor, so we have only to specify the spins just outside the boundary. However, if the interaction is wider-ranged, we must specify a lot about the external 'spin configurations.'

[8] Here, $V$ denotes a bounded (and singly connected) subset of the whole lattice, but occasionally also denotes its volume.

limit points (= probability measures) available through this construction (that is, accumulation points of $M$) are called the *Gibbs measures* (however, read the fine-lettered explanation on the "totality of Gibbs measure" below).

To speak about 'convergence' we must specify the topology of the space $M$. That is, we must be able to tell which canonical distribution of the form (5.2.1) is close to which. Since our aim is statistical mechanics, it is sensible to define the closeness of two distributions by the closeness of thermodynamic observables computed by these distributions. Therefore, we interpret the statement "The probability measure sequence $\{\mu_i\}$ converges to $\mu$," as "all the thermodynamic observables computed by this sequence converge to those computed by $\mu$."

**At least one Gibbs state exists**
An argument goes like...[9]
Here, we mean by a lattice system a system consisting of discrete entities fixed at lattice points that can take (at most) countably many states. The existence of a convergent sequence in $\{\mu_{V,B}\}$ is shown by the *diagonal argument*. We take a countable list of bounded thermodynamic observables $\{Q_i\}$ (i.e., we make a list of observables such as internal energy density, magnetization density, etc.) Let us write the expectation value of observable $Q$ with respect to $\mu$ as $\mu(Q)$. Let us first make a sequence $\{\mu_{V,B}(Q_1)\}$ of observation results of $Q_1$. Since this is a bounded sequence, we can choose a converging sequence $\{\mu_{1i}(Q_1)\}$. Next, we make the sequence $\{\mu_{1i}(Q_2)\}$ for observable $Q_2$, and choose a convergent subsequence $\{\mu_{2i}(Q_2)\}$. Repeat this procedure to make $\{\mu_{ki}\}$ for all $k \in \boldsymbol{N}$. Now, choose the diagonal elements and make a sequence $\{\mu_{kk}\}$. By construction $\{\mu_{kk}(Q_i)\}$ converges for any $Q_i$. Thus we may conclude that $\{\mu_{kk}\}$ converges to a $\mu$, which is a Gibbs measure. Thus, there is at least one Gibbs measure for any reasonable lattice model.

A possible definition of the phase transition for a system is the occurrence of qualitative change in the set of the totality of the Gibbs states of the system. For example, if the temperature is lowered, the number of (pure) Gibbs states changes 1 to 2 for 2-Ising model (see Sect. 5.4). This temperature is the critical point = the second order phase transition point.

**The totality of Gibbs states**
In the text thermodynamic limit is understood as a limit reached by a sequence of increasingly large systems, and the Gibbs measure is defined as a measure obtained in this sequence limit. However, in mathematical physics, the Gibbs measure is directly defined without any limiting procedure (classically by the DLR (Dobrushin-Lanford-Ruelle) equation, quantum-mechanically by the KMS (Kubo-Martin-Schwinger) con-

---

[9] That $\{\mu_{V,B}\}$ is a compact set with respect to the vague topology is perhaps the best argument for mathematicians. The argument here with the aid of the diagonal argument is quite general, but sounds rather artificial. For general systems, more physical arguments would not be available.

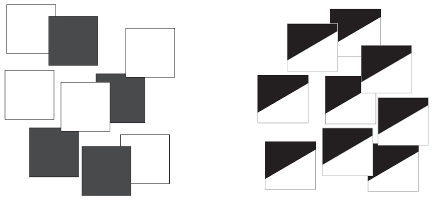
dition, for example).[10] The totality $K$ of the Gibbs measures defined directly is the convex linear hull of the Gibbs measures defined sequentially. $K$ is a nonempty convex set and any state in it must be one of the following (a) or (b):[11]

(a) $\mu \in K$ cannot be expressed as a linear combination of other members of $K$. Such states are called *pure states*.**r** (**r**to symmetry breaking)

(b) $\mu \in K$ can be uniquely decomposed into a linear combination of pure states. Such states are called *mixed states*.

Physically, pure states correspond to macroscopic states that can be observed locally for a sufficiently large single system.[12] Do not confuse pure Gibbs states and thermodynamic single phases (pure phases). A mixed state corresponds, as illustrated in Fig. 5.1.3, to an ensemble containing different (pure) states and has nothing to do with individual samples to have more than one phases.[13]

The macroscopic thermodynamic state we observe in a single system corresponds to a pure Gibbs state. Precisely speaking, a pure Gibbs state corresponds to a unique thermodynamic state of the system. However, a thermodynamic state is defined as the equivalence class with respect to the thermodynamic coordinate values, so it corresponds to an equivalence class of pure Gibbs states with respect to the thermodynamic coordinate values. For example, if two phases coexist in a pure Gibbs state, the phase boundary plane is also specified in space; however, thermodynamic states are indifferent to the actual (normal) direction of the phase boundary.



mixture ensemble

pure state ensemble

Fig. 5.1.3 Two kinds of ensemble; a pure Gibbs state corresponds to an actual state of a single sample observable by a single observation. Notice that both ensembles give the same expectation values for thermodynamic quantities. In this illustration the left figure exhibits a mixture Gibbs state made of thermodynamically pure phase samples, and the right figure exhibits a pure Gibbs states whose samples contain thermodynamically coexisting phases.

We have reached an idea that the phase transition is a point where the Gibbs state of the system changes qualitatively. It is also natural to expect that at such a point thermodynamic functions should have certain qualitative changes (have singularities). Let us consider the singular points of thermodynamic functions.

---

[10] A good reference is R. B. Israel, *Convexity in the Theory of Lattice Gases* (Introduction by A. S. Wightman) (Princeton Univ. Press 1979). The introduction by Wightman should be read by everyone.

[11] That is, $K$ is a Choquet simplex.

[12] Here, 'locally' implies that we observe a large but finite volume at around, e.g., the origin (or in front of us).

[13] Sequentially defined limits need not be pure states.

## 5.3 Phase transition and singularity

rIn this section let us observe the relation between the phase transition and the singularity of a thermodynamic potential (e.g., free energy loses differentiability at the phase transition point) with the aid of a classical fluid consisting of particles with sufficiently hard cores.

If the system is finite (e.g., with a volume $V$), then the grand canonical partition function reads

$$\Xi_V(z) = \sum_{N=0}^{M} z^N Z_N(T,V), \tag{5.3.1}$$

where $M$ is the upper bound of particles pushed into $V$,[14] $z = e^{\beta\mu}$ (= fugacity) and

$$Z_N(T,V) = \frac{1}{h^{3N}N!}\int d\Gamma_N \, e^{-\beta H_N} \tag{5.3.2}$$

may be understood as the canonical partition function for a $N$ particle system. Also we assume that the potential energy part $U_N$ of the system Hamiltonian $H_N$ when there are $N$ particles in $V$ is bounded from below as

$$U_N \geq -NB, \tag{5.3.3}$$

with some positive number $B$. (5.3.1) is a polynomial of $z$, so it is obviously holomorphic[15] for any complex $z$ (i.e., $\Xi_V(z)$ is entire). The pressure of the system may be obtained as

$$\frac{P_V}{k_BT} = \frac{1}{V}\log \Xi_V(z). \tag{5.3.4}$$

Here, suffix $_V$ explicitly denotes the finiteness of the system. Since all the coefficients of the polynomial $\Xi_V(z)$ are positive, it cannot have any real positive zero. Therefore, $P_V$ is holomorphic on the real positive axis. Now, it is clear that there is no ('clear') phase transition for finite systems. The derivative

$$\frac{\partial}{\partial \log z}\beta P_V = n_V > 0 \tag{5.3.5}$$

---

[14] Here, for simplicity, let us assume that the system, if finite, is confined in an infinite potential well.

[15] i.e., locally expandable into Taylor series; physicists often say 'analytic,' intead. Mathematically, 'analyticity' is a global concept. In contrast, 'holomorphy' is a local concept.
⟪**Holomorphy**⟫ A complex valued function defined on $\boldsymbol{C}$ is *holomorphic* on an open set $\Omega$, if its Taylor expansion around any point in $\Omega$ has a finite convergence radius. [If $\Omega$ is not an open set, we interpret that holomorphy holds in an appropriate neighborhood of $\Omega$.]

is the number density, which is bounded by the densest close packing density, so it is uniformly bounded for all $V$. Therefore, if the thermodynamic limit exists, then the limit of this derivative is bounded. That is, the limit $P$ of $P_V$, if exists in the $V \to \infty$ limit, is continuous and monotone increasing function of $z$.

The proof of the existence of thermodynamic limit consists of demonstrating of the following three steps:
(1) In the large volume limit, the surface effect can be ignored for a cube.
(2) If $\{V\}$ is a nested increasing sequence of cubes, then $\{(1/V)\log \Xi_V(z)\}$ converges.[16]
(3) For any nested sequence of volumes of any shape satisfying the van Hove condition, the large volume limit converges.

**Existence of thermodynamic limit of $P$**
For physicists (1) should be clear. (3) should not be a problem, either, if we admit (2) (see Fig. 5.3.1).

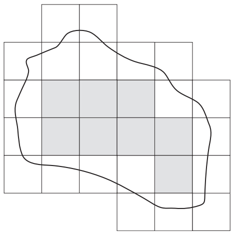


Fig. 5.3.1 The thermodynamic limit for a general domain, but with not too bad surface geometry. The domain is tessellated with the cubes from inside and from outside. In the thermodynamic limit, the majority of the cubes do not intersect the boundary.

(2) may be demonstrated as follows:

Let $W$ be (the volume of) a cube with a given edge length. Let us subdivide $W$ into $n$ subcubes, and write the grand canonical partition function for a subcube as $\Xi_{W/n}$. The number of particles in the boundary zone may be bounded from above by a number proportional to $6(W/n)^{2/3}n \propto n^{1/3}W^{2/3}$ (the total surface area of the subcubes). Therefore, the energetic contribution of these boundary particles is bounded from below by $-k_BT\gamma n^{1/3}W^{2/3}$,[17] where $\gamma > 0$, because the interaction is attractive. Consequently, we have

$$\Xi_W \le \Xi_{W/n}^n e^{n^{1/3}W^{2/3}\gamma}. \tag{5.3.6}$$

This implies that

$$\log \Xi_W \le n \log \Xi_{W/n} + n^{1/3}W^{2/3}\gamma. \tag{5.3.7}$$

In short, if we write $f(V) = \log \Xi_V$, this function is subadditive (up to a higher order correction). Therefore,[18] $\log \Xi_V/V$ converges to $\inf_V (1/V)\log \Xi_V$. $\Xi_V > 1$ and it is

---

[16] Here, the same symbol $V$ is also used to denote its volume.

[17] Here, we have assumed that the system is stable.

[18] due to Fekete's lemma:

also bounded from above by $M$ times the largest summand which is bounded from above by $e^{cV}$ for some positive $c$, so this limit is finite.

The pressure $P_V$ is bounded for finite $z$ from above for all $V$ uniformly as there is a constant such that $\Xi_V < e^{cV}$ as can be shown with the aid of (5.3.3). Therefore, Vitali-Porter's theorem tells us that in the domain $\mathcal{R}$ where no zero of $\Xi_V$ appears for all sufficiently large $V$, $e^{\beta P}$ is holomorphic. Hurwitz' theorem tells us that its zero is only the accumulation point of the zeros of $\Xi_V$.

**Needed elementary theorems from complex analysis**
For convenience, theorems of complex analysis used in the text are summarized here; it is a good occasion to review complex analysis.[19]
**Theorem** [Vitali (1903), Porter(1904)] Let $D$ be a region (= open connected set in $\boldsymbol{C}$) on which a sequence $\{f_n\}$ is holomorphic. If the sequence is (locally) uniformly bounded, and is convergent on a set that have an accumulation point in $D$, then $f_n$ converges (locally) uniformly in $D$.
This is often an exercise question related to the Vitali-Montel theorem.[20]

**Theorem** [Hurwitz] Let $\{f_n\}$ be a sequence of holomorphic functions in a region $D$ with $\partial D$ being a simple closed curve. If the convergence $f_n \to f$ is uniform in $D$ and $f$ is not identically zero there ($f(z) \not\equiv 0$), then a necessary and sufficient condition for $z_0$ to be a zero of $f$ in $D$ is that $z_0$ is an accumulation point of the zeros of $f_n(z)$.
This follows immediately from the argument principle.[21]

We thus conclude:[q]
(1) As $V$ grows the number of zeros of $\Xi_V$ increases. If they do not come near the real positive axis, still we would not have any phase transition. However, if they close in onto a certain part of the real positive axis, then the holomorphy of $\lim_V P_V(z)$ is lost there, and we must conclude that there is a phase transition. The distribution of zeros depends on temperature, so the fugacity (or the density) at which phase transition occurs depends on temperature.

---

⟪**Fekete's lemma**⟫ If $f$ is subadditive (i.e., $f(n+m) \le f(m) + f(n)$ for any $n, m \in \boldsymbol{N}$), then $\lim_{n\to\infty} f(n)/n = \inf f(m)/m$.
[Demo] Obviously, $\liminf f(n)/n \ge \inf f(m)/m$. Let $n = s + km$. Then,

$$\frac{f(n)}{n} = \frac{f(s+km)}{n} \le \frac{f(s)+kf(m)}{s+km} \to \frac{f(m)}{m}.$$

Therefore, $\limsup f(n)/n \le \inf f(m)/m$.

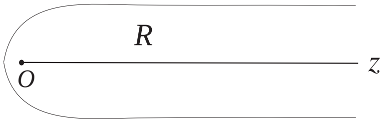
[19] A good and modern reference of complex analysis may be M. Rao and H. Stetkaer, *Complex Analysis, an invitation* (World Scientific, 1991).

[20] **Theorem** [Vitali-Montel] Let $\Omega$ be open in $\boldsymbol{C}$, and $\{f_k\}$ be a sequence of holomorphic functions on $\Omega$ that are uniformly bounded on any compact subset in $\Omega$. Then, there is a uniformly convergent subsequence on every compact subset of $\Omega$.

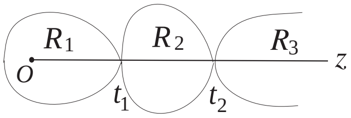
[21] $\int_{\partial D}(f'/f)dz$ is the number of zeros in $D$, if $f$ is holomorphic in $D$ and not zero on $\partial D$.

(2) We have already seen that the virial expansion of the pressure has a positive convergence radius,[22] so if the fugacity is small enough, $P$ is holomorphic in $z$. That is, there is no phase transition if $\beta\mu$ is small (high temperature and/or low density). We may assume that it is a thermally agitated disordered phase. It is unique for any reasonable system.

This is the picture first explicitly recognized by Lee and Yang.[23]



no phase transition



two phase transitions

Fig. 5.3.2 The zero free regions on the complex fugacity plane and phase transitions. If, for example, zeros of $\Xi_V(z)$ close in onto the real positive axis in the thermodynamic limit at $t_1$ and $t_2$, then the zero free region is decomposed into three regions. The interpretation is that each phase (say, gas, liquid, etc) corresponds to each region, and that phase transitions occur at the fugacities where zeros close in.

Thus, a phase transition may be mathematically characterized by a singularity of a thermodynamic potential. The singularity here implies the loss of holomorphy, so the potential may still be $C^\infty$ even at the phase transition.[24] This book adopts this definition of phase transition. The bifurcation in the Gibbs states certainly implies phase transition, but the converse may not hold.

Then, what is a *phase*?**r** We may be tempted to say that states in, say, $R_2$ in Fig. 5.3.2 all belong to the same phase; a phase corresponds to an analytic branch of a thermodynamic potential. However, even in this case, if we change the temperature $R_2$ and $R_3$ may merge. This is just what happens between the gas and the liquid phases. We can go to one phase from the other without any phase transition. Even if we may not find such a path in the ordinary phase diagram, perhaps we can apply an exotic field to connect two 'phases' continuously by changing the field. Therefore, in this book 'phase' is understood as a local concept relative to a phase transition: if

---

[22] For the lattice gas model below this is guaranteed by the Lee-Yang circle theorem.

[23] C. N. Yang and T. D. Lee, Phys. Rev. **87** 404 (1952), "Statistical theory of equations of state and phase transitions. I. Theory of condensation." This is a classic definitely worth reading; Proofs can be made simple with the aid of standard theorems, but we realize that it is a sign of good theoreticians that everything needed are proved without known theorems.

[24] Such examples indeed exists: the 2-dimensional XY-model.

we have a phase transition, then near the phase transition point we can clearly say phases are distinct.

It is not easy to demonstrate the above general picture for realistic models. Lee and Yang illustrated the general theory with the aid of a solvable model.[25] To understand this first let us discuss the *Ising-lattice gas correspondence*.**q** We interpret down spins of the Ising model as particles, and up spins as vacancies. In this way the spin system is translated into a lattice gas system. The *lattice gas model* consists of a lattice whose individual lattice points may be occupied by at most one particle. Let $n_i \in \{0, 1\}$ be the occupation number of lattice point $i$. Then, the model Hamiltonian reads

$$H_L = -J' \sum_{\langle i,j \rangle} n_i n_j, \tag{5.3.8}$$

where $J' > 0$ is a coupling constant (nearest neighbor particles attract each other).

Let us consider an Ising model on a cubic lattice. The canonical partition function[26] for the Ising model reads**q**

$$Z = \sum_{SC} e^{\beta J \sum_{\langle i,j \rangle} s_i s_j + \beta h \sum_j s_j}, \tag{5.3.9}$$

where $\sum_{SC}$ is the summation over all the spin configurations. Let $[D]$ be the number of down spins, and $[U]$ up spins. Then

$$\sum_i s_i = -[D] + [U] = V - 2[D], \tag{5.3.10}$$

where $V = [D] + [U]$ is the total number of lattice points (= volume). Let $[UD]$ be the total number of the up-down (antiparallel) spin pairs. Then,

$$\sum_{\langle i,j \rangle} s_i s_j = 3V - 2[UD], \tag{5.3.11}$$

so we may write the total energy as

$$-J \sum_{\langle i,j \rangle} s_i s_j - h \sum_j s_j = -3JV + 2J[UD] + h(2[D] - V). \tag{5.3.12}$$

---

[25] The following is an introduction to T. D. Lee and C. N. Yang, "Theory of condensation, II. Lattice gas and Ising model," Phys. Rev., **87**, 410-419 (1952).

[26] It is under an external magnetic field, so it is a generalizd canonical partition funcition, but we call it simply 'canonical.'

Let $[DD]$ be the total number of down spin pairs. There are two ways to count the number of 'arms' sticking out from all down spins: there are 6 arms from each down spin, and one down spin pair uses 2 arms and one up-down spin pair 1 arm. Therefore,[27]

$$2[DD] + [UD] = 6[D]. \tag{5.3.13}$$

Therefore, (5.3.12) reads

$$-J\sum_{\langle i,j\rangle} s_i s_j - h\sum_j s_j = -3JV + 12J[D] - 4J[DD] + h(2[D] - V). \tag{5.3.14}$$

Consequently, the canonical partition function of the Ising model may be rewritten as

$$Z = e^{-\beta V f} = \sum_{SC} \exp\left[-\beta\left(-(3J+h)V + (2h+12J)[D] - 4J[DD]\right)\right]. \tag{5.3.15}$$

Here, $f$ is the free energy per spin. The formula implies that the canonical partition function of the Ising model can be translated into the grand canonical partition function of the lattice gas model (note that $H_L = -J']DD]$):

$$e^{\beta PV} = \sum_{SC} z^{[D]} e^{4\beta J[DD]} = \sum_{N=0}^{V} z^N \sum_{PC,N} e^{-\beta H_L}. \tag{5.3.16}$$

Here, $\sum_{PC,N}$ implies the sum over all the $N$ particle configurations, the fugacity is $z = e^{-\beta(2h+12J)}$, $N = [D]$, the pressure of the gas is $P = -(3J + h + f)$ and $H_L$ is (5.3.8) with $J' = 4J$.[28]

If (5.3.16) is considered as a polynomial of $z$, its roots are all on a circle ($|z| = e^{-12\beta J}$) according to the circle theorem below. This implies that there is at most one phase transition for the lattice gas when the fugacity is changed. The most general form of the *circle theorem* is as follows:

**Lee-Yang Circle theorem**$_{\mathbf{q}}$ Let the real numbers $a(\{i,j\}) \in [-1,1]$ for all $\{i,j\}$ ($i,j \in L = \{1,\cdots,V\}$). Then, all the zeros of the following polynomial are on the unit circle:

$$P(y) = \sum_{X\subset L} y^{|X|} \prod_{i\in X}\prod_{j\notin X} a(\{i,j\}). \tag{5.3.17}$$

---

[27] Notice that this is exact only on the infinite square lattice; if the lattice is very small (or the portion of spins near the lattice edges cannot be ignored) the magnet-lattice gas translation here is not applicable.

[28] This correspondence is correct in the large lattice limit.

Here, the constant term and the coefficient of the highest order term $y^V$ must be unity and $|X|$ is the number of elements in $X$ (the cardinality of $X$). The summation is over all the subsets of $L$. $\square$[29]

If we set $y = ze^{12\beta J} = e^{-2\beta h}$, as a polynomial of $y$ (5.3.16) is $P(y) = \sum y^{[D]} e^{-2\beta J[UD]}$, satisfying the conditions in the circle theorem. The theorem asserts that for any regular or random lattice and for any finite range (random) ferromagnetic coupling among spins, the corresponding lattice gas has at most one phase transition when the fugacity (or pressure) is changed. However, its assertion about the original magnetic system is uninteresting. The theorem does not tell us anything about the phase transition due to temperature change. When magnetic field is changed, there is at most one phase transition and that occurs at $y = 1$, i.e., $h = 0$.

The correspondence between the Ising model and the lattice gas model suggests that there are common features in the phase diagrams of a magnet and of a fluid system. Furthermore, we may interpret the Ising model as a lattice fluid mixture of 'up' molecules and 'down' molecules (or the fluid as a mixture of molecules and vacancies), so the phase diagram of a binary mixture must share some features with that of magnets (Fig. 5.3.3).

## 5.4 Phase transition and fluctuation

rPhase ordering is possible because the order can resist thermal fluctuation.(rto crystal) To this end microscopic entities must stand 'arm in arm.' The number of entities each entity directly interacts (cooperates) crucially depends on the spatial dimensionality. Let us look at the effect of spatial dimensionality on the Ising model.

**1-Ising model**:
We can obtain the free energy (with magnetic field) exactly as we will see later by, e.g., the transfer matrix method (see Section 5.9); the phase transition does not occur for $T > 0$ as we have intuitively seen in Section 5.1 and will see soon by calculation that $T = 0$ is the critical point.

**2-Ising model**:r

---

[29] The most succinct proof is Appendix to D. Ruelle, "Is our mathematics natural? The case of equilibrium statistical mechanics," Bull. Amer. Math. Soc., **19**, 259 (1988). An accessible version is in Supplement Pages.

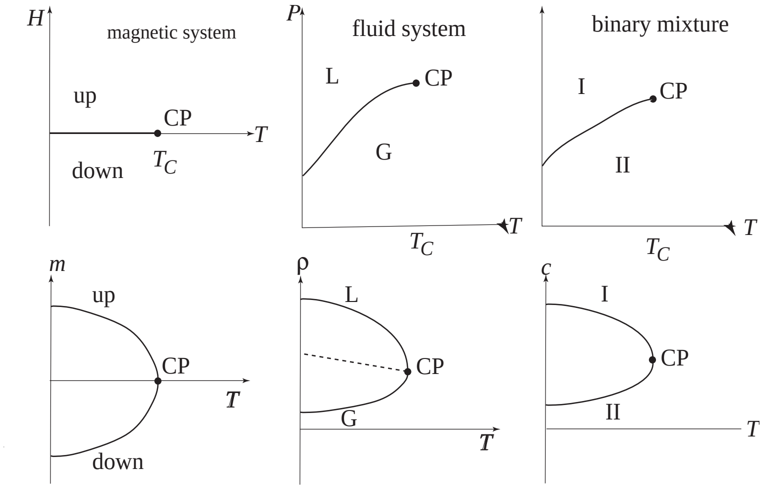


Fig. 5.3.3 The correspondence between the magnetic, the fluid, and the binary liquid mixture systems near the critical point CP. $T$: temperature, $T_c$: the critical temperature, $H$: magnetic field, $P$: pressure, $\mu$: chemical potential of one component, $m$: magnetization per spin, $\rho$: the number density, $c$: the concentration of a particular component.

For the magnetic system, the spins are assumed to be the Ising spins (only two directions are allowed, up or down), and 'up' (resp., 'down') in the figure means majority of the spins point upward (resp., downward) (ferromagnetically ordered). L implies the liquid and G the gas phase. I and II denote different mixture phases.

The following correspondences are natural: for the fields $H \leftrightarrow P \leftrightarrow \mu$; for the order parameters $m \leftrightarrow (\rho_L - \rho_G) \leftrightarrow (c_I - c_{II})$.

(1) The Onsager solution gives the free energy without magnetic field.[30] There is a phase transition at $T_c > 0$.
(2) Below the phase transition temperature $T_c$ there are only two pure Gibbs states corresponding to the up spin phase and the down spin phase.[31] That is, up and down phases cannot coexist (see Fig. 5.4.1).r

---

[30] L. Onsager, "Crystal statistics. I. A two-dimensional model with an order-disorder transition," Phys. Rev., **65**, 117-149 (1944). Probably, the easiest method is to use Grassmann algebra: S. Samuel, "The use of anticommuting variable integrals in statistical mechanics. I. The computation of partition functions," J. Math. Phys., **21**, 2806 (1980). This is explained in Supplementary Page.

[31] Independently by: M. Aizenman, "Translation invariance and instability of phase coexistence in the two-dimensional Ising model," Commun. Math. Phys. **73**, 83 (1980); Y. Higuchi, "On the absence of non-translationally invariant Gibbs states for the two-dimensional Ising system," in *Random Field* (ed. J. Fritz, J. L. Lebowitz and D. Szaz, North-Holland, 1981).

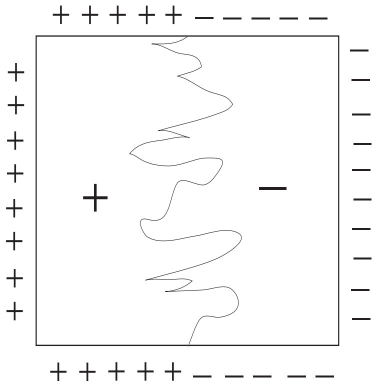


Fig. 5.4.1 Even the half up and half down fixed boundary spin configuration cannot stabilize the interface location between up and down phases for 2D Ising model below $T_c$. The interface may be understood as a trajectory of a Brownian particle connecting the two phase boundary points at the boundary (Brownian bridge). If the system size is $L$, then its amplitude is $\sqrt{L}$. In the thermodynamic limit almost surely the observer at a fixed point (say, at the center) who can observe only a finite volume can observe only one of the phases, and can never see the spin flip in her lifetime.

(3) The effect of the magnetic field can be calculated perturbatively, so magnetic susceptibility is exactly known.[32],[33]
(4) Near $T_c$ there are various nontrivial *critical divergences*.[34]

**3-Ising model**:
(1) No exact evaluation of the free energy is known, but it is easy to demonstrate that $T_c > 0$ (see Peierls' argument). It is known that at sufficiently low temperatures there are infinitely many pure Gibbs states.[35] They are interpreted as up and down spin phases and their coexisting phases.
(2) Thus, at sufficiently low temperatures the interface is not rough. Whether this happens at $T_c$ (where the uniqueness of the Gibbs state is lost) or strictly below it (that is, there is a separate *roughening transition*) cannot be known from the existing theories, but according to simulations the roughening transition on the 001 surface of the cubic Ising model occurs about $T_c/2$.[36]

---

[32] C. N. Yang, "The spontaneous magnetization of a two-dimensional Ising model," Phys. Rev. **85**, 808 (1953).

[33] In the scaling limit, the system with a magnetic field can be solved exactly (due to Zamolodchikov). See G. Delfino, "Integrable field theory and critical phenomena: the Ising model in a magnetic field," J. Phys. A **37** R45 (2004).

[34] Here, 'non-trivial' means that the fluctuation is so large that we cannot use mean-field theory to study the divergent behavior correctly.

[35] R. L. Dobrushin, "Gibbs state describing coexistence of phases for a three-dimensional Ising model," Theor. Probab. Applic., **17**, 582 (1972); "Investigation of Gibbsian state for three-dimensional lattice systems," ibid., **18**, 253 (1973).

[36] See M. Hasenbusch and K. Pinn, "Computing the roughening transition of Ising and solid-on-solid models by BCSOS model matching," J. Phys. A **30**, 63 (1997).

(3) The critical divergences are non-trivial as we will see.

**Beyond 3-space**:
Although no exact free energy is known, the existence of positive $T_c$ is easy to demonstrate, and the critical divergences around this point are believed to be the same for all $d \geq 4$. This has been established for the dimension strictly greater than 4;[37] 4-Ising model still defies mathematical studies.

Phase transitions are often studied by changing intensive parameters (e.g., temperature and pressure). When two phases coexist, they share the same intensive parameters (fields). Therefore, a convenient thermodynamic potential is the *generalized Gibbs free energy* for a given amount of material ($N$); for example, that obtained by Legendre transformation of internal energy with respect to entropy, volume, magnetization, etc., except for the number of particles $N$. The Gibbs free energy may lose differentiability with respect to its natural independent variables (intensive parameters). If the differentiability is lost, we say a *first order phase transition* occurs. If the singularity in Gibbs free energy is less drastic, generally we say there is a *second order phase transition*.[38]

The reader might conclude that the first order phase transition is the most basic phase transition, but first order phase transition occurs when an ordered phase becomes unstable against not very large thermal fluctuations. Therefore, we could say the ordered phase that disappears by a first order phase transition is not so stable.**r** In the case of the second order phase transition the ordered phase withstands thermal fluctuation as long as there is 'positive order' (the order parameter is positive). The 'ultimate phase transition' should be the disappearance of 'very stable' order due to very large fluctuations. Thus, the second order phase transition is the most typical phase transition. In equilibrium states, fluctuations cannot become large without long-range correlation. Consequently, if the ordered phase is destroyed only when the fluctuation is very large (divergent, mathematically), the range of the correlation of order parameter must be unbounded. It must be crucial to study the correlation of order parameters to understand the second-order phase transition.

---

[37] M. Aizenman, "Proof of the triviality of $\varphi^4_d$ field theory and some mean-field features of Ising models for $d > 4$, Phys. Rev. Lett., **47**, 1 (1981).

[38] The terminology, "infinite order phase transition" is used if the free Gibbs energy loses holomorphy (i.e., not $C^\omega$) but still infinite times differentiable ($C^\infty$). This can happen in 2D XY model.

## 5.5 Correlation of oder parameter

For second order phase transitions fluctuations are crucial. To see this let us calculate the order parameter correlation function and study the *correlation length* $\xi$ defined as follows:[39]

$$\langle \phi(0)\phi(\boldsymbol{r})\rangle \sim e^{-|\boldsymbol{r}|/\xi}. \tag{5.5.1}$$

Here, $\sim$ implies a long distance asymptotic relation.[40]

At the second order phase transition point, the correlation length diverges. What happens to the correlation function? (5.5.1) implies that the correlation must decay slower than any exponential decay. For some $c > 0$ the decay like $e^{-c\sqrt{r}}$ can be imagined, but such a stretched exponential decay is impossible for equilibrium spatial correlation functions. For example, when ferromagnetic ordering occurs, it can be proved[41] that

$$\langle \phi(0)\phi(\boldsymbol{r})\rangle \le \sum_{\boldsymbol{s}\in B}\langle \phi(0)\phi(\boldsymbol{s})\rangle\langle \phi(\boldsymbol{s})\phi(\boldsymbol{r})\rangle. \tag{5.5.2}$$

Here, $B$ can be any set of lattice points such that removal of all the points in $B$ destroys all the paths along the lattice bonds connecting the origin 0 and $\boldsymbol{r}$. For example, we may choose it to be a spherical shell centered at the origin with an appropriate thickness.[42] If the system is translationally symmetric, then from (5.5.2) we get

$$\langle \phi(\boldsymbol{s})\phi(\boldsymbol{r})\rangle = \langle \phi(0)\phi(\boldsymbol{r}-\boldsymbol{s})\rangle \le \sum_{\boldsymbol{s}_1\in B}\langle \phi(0)\phi(\boldsymbol{s}_1)\rangle\langle \phi(\boldsymbol{s}_1)\phi(\boldsymbol{r}-\boldsymbol{s})\rangle \tag{5.5.3}$$

All the moments of spins are positive.[43] Therefore, introducing the relation (5.5.3) into (5.5.2), we obtain

$$\langle \phi(0)\phi(\boldsymbol{r})\rangle \le \sum_{\boldsymbol{s}\in B}\langle \phi(0)\phi(\boldsymbol{s})\rangle\left(\sum_{\boldsymbol{s}_1\in B}\langle \phi(0)\phi(\boldsymbol{s}_1)\rangle\langle \phi(\boldsymbol{s}_1)\phi(\boldsymbol{r}-\boldsymbol{s})\rangle\right). \tag{5.5.4}$$

---

[39] Here, we assume $\langle\phi\rangle = 0$. More generally, we should use the second order cumulant: $\langle \phi(0)\phi(\boldsymbol{r})\rangle_C = \langle \phi(0)\phi(\boldsymbol{r})\rangle - \langle \phi(0)\rangle\langle \phi(\boldsymbol{r})\rangle$.

[40] Precisely, the ratio of the logarithms of the both sides reaches unity in the $|\boldsymbol{r}| \to \infty$ limit.

[41] For ferromagnetic Ising models, see B. Simon, "Correlation Inequalities and the decay of correlations in ferromagnets," Commun. Math. Phys., **77**, 111 (1980).

[42] For $d$-(hyper)cubic lattice, its thickness can be slightly larger than $\sqrt{d}\times$ lattice spacing.

[43] This is one of *Griffiths' inequalities* (the first inequality): for any finitely many positive integers $a_i$ $\langle\prod_i s_i^{a_i}\rangle \ge 0$. A simple proof may be found in J. Glimm and A. Jaffe, *Quantum Physics, a functional integral point of view*, Second Edition (Springer, 1987) Sect. 4.1.

To be sure let us reiterate this procedure once more:

$$\langle\phi(0)\phi(\boldsymbol{r})\rangle \le \sum_{\boldsymbol{s}\in B}\langle\phi(0)\phi(\boldsymbol{s})\rangle\left\{\sum_{\boldsymbol{s}_1\in B}\langle\phi(0)\phi(\boldsymbol{s}_1)\rangle\left(\sum_{\boldsymbol{s}_2\in B}\langle\phi(0)\phi(\boldsymbol{s}_2)\rangle\langle\phi(\boldsymbol{s}_2)\phi(\boldsymbol{r}-\boldsymbol{s}-\boldsymbol{s}_1)\rangle\right)\right\}. \tag{5.5.5}$$

If $\boldsymbol{r}$ is sufficiently far away from the origin and the above procedure may be repeated $q$ times, then replacing the term $\langle\phi(\boldsymbol{s}_q)\phi(\boldsymbol{r}-\boldsymbol{s}-\boldsymbol{s}_1-\cdots-\boldsymbol{s}_{q-1})\rangle$ that is expected to appear after these iterations with its maximum value $C$, we obtain

$$\langle\phi(0)\phi(\boldsymbol{r})\rangle \le \left[\sum_{\boldsymbol{s}\in B}\langle\phi(0)\phi(\boldsymbol{s})\rangle\right]^q C. \tag{5.5.6}$$

Here, $q$ may be chosen to be a number proportional to $|\boldsymbol{r}|$. If we can make $\sum_{\boldsymbol{s}\in B}\langle\phi(0)\phi(\boldsymbol{s})\rangle < 1$, the correlation function decays exponentially. That is, the summability of the correlation function on a large sphere is the key for exponential decay. If the correlation function behaves as $r^{-\mu}$ $(d-1<\mu)$, the decay must be exponential. On the other hand, as can be seen from the decay of the fundamental solution to the Laplace equation, even without any fluctuation, correlation decays as $1/r^{d-2}$.

If there are fluctuations, intuitively speaking at least for the ferromagnetic case, the correlation is expected to decay faster than without fluctuations. Therefore, when $\xi$ diverges, the correlation function must decay algebraically; if we write this algebraic decay as $r^{-\mu}$, $\mu$ cannot be larger than $d-1$, and cannot be smaller than $d-2$. Therefore, we may conclude that[r]

$$\langle\phi(0)\phi(\boldsymbol{r})\rangle \sim \frac{1}{r^{d-2+\eta}} \tag{5.5.7}$$

is the general expression for the order parameter correlation function at the critical point. Here, $\eta$ $(\in(0,1])$ is one of the critical indices we will encounter in the next section.

The above argument suggests that if there is a long range interaction, fluctuation effects can be contained. We already know such an example: the 1D Kac model. Even in 1-space this long-range interaction model exhibits gas-liquid phase transition.[44]

What happens if we introduce a small number of (fraction of) long range interactions into a short-ranged lattice system? This is a story of systems with the so-called

---

[44] However, no local order can be stabilized by a long range interaction, so no crystal formation is possible.

'small world' interaction network.[45] Let us introduce such long range interaction with probability $p$ ($\ll 1$). The average distance from an arbitrary lattice point to one of such long-range interacting points (= airports) is $\ell \sim p^{-1/d}$. We compare this with the domain size = correlation length $\xi$, which is the distance over which spins can 'communicate.' This length generally grows as temperature decreases. The ordering would occur if $\xi$ reaches $\ell$, because then the ordered patch can have access to the 'big world.' As we will learn, in 1-space[46]

$$\xi \sim e^{2J/k_BT}, \tag{5.5.8}$$

so $T_c \sim 1/|\log p|$; now phase transition can occur at a positive temperature. Also as we will learn later $\xi \sim (T - T_c)^{-\nu}$ in $D(>1)$-space, so $T_c(p) - T_c(0) \sim p^{1/d\nu}$.

## 5.6 Critical fluctuation and Kadanoff picture

Second order transitions**r** are characterized by the divergence of correlation length $\xi$ associated with very large equilibrium fluctuations:**q**

$$\xi \sim |T - T_c|^{-} = |\tau|^{-\nu}, \mathbf{r} \tag{5.6.1}$$

where $\tau = (T - T_c)/T_c$, and $\nu$ ($> 0$) is an example of *critical exponents* (*critical indices*). We know susceptibilities directly reflect fluctuations as is clearly seen from a fluctuation-response relation:

$$\langle \delta M \delta M \rangle = k_BT \left.\frac{\partial M}{\partial H}\right|_T. \tag{5.6.2}$$

This implies that at the second order phase transition point the magnetic susceptibility $\chi$ diverges. Indeed, empirically it diverges as (for $h = 0$, without magnetic field)

$$\chi \sim |T - T_c|^{-\gamma} = |\tau|^{-\gamma} \ (h = 0). \mathbf{r} \tag{5.6.3}$$

---

[45] see D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton Studies in Complexity, 2003).

[46] The probability of introducing one up-down boundary is $e^{-2J/k_BT}$, so the spacing of 'defects' should be the reciprocal of this.

We cannot expect smooth change of the magnetization $m$,[47] so

$$m \sim (-\tau)^{\beta} \ (h=0, \tau<0), \ \mathbf{r} \tag{5.6.4}$$

$$m \sim h^{1/\delta} \ (\tau=0).\mathbf{r} \tag{5.6.5}$$

The divergence of energy fluctuation causes the divergence of specific heat as

$$C_H \sim |\tau|^{-\alpha} \ (h=0).\mathbf{r} \tag{5.6.6}$$

$\alpha$, $\beta$, $\gamma$, $\delta$ are positive numbers and are critical indices. Representative values can be found in the following table. It was empirically noted that several relations hold among these indices such as

$$\alpha + 2\beta + \gamma = 2. \tag{5.6.7}$$

Ising critical exponents.**r**

| Exponents | 2-space | 3-space | $d(\geq 4)$-space |
|---|---|---|---|
| $\alpha$ | 0 (log) | 0.11 | 0 (jump) |
| $\beta$ | 1/8 | 0.325 | 1/2 |
| $\gamma$ | 1.75 | 1.24 | 1 |
| $\delta$ | 15 | 4.8 | 3 |
| $\nu$ | 1 | 0.63 | 1/2 |

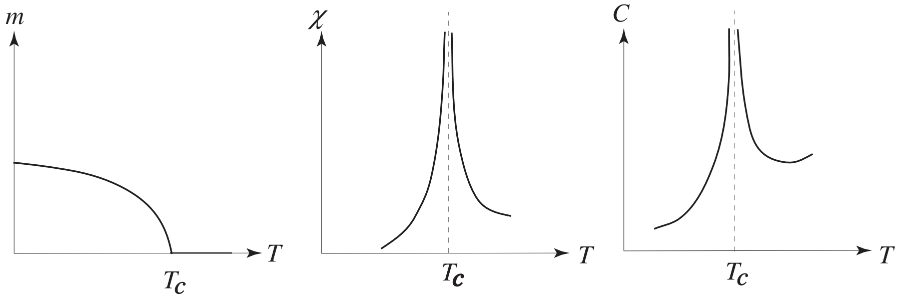


Fig 5.6.1 Schematic illustrations of singular behaviors near the critical point.

A typical spin configuration at the critical point of the Ising model on the square lattice looks as in Fig. 5.6.2.[48]

---

[47] Notice that there is no logical relation between the divergence of the susceptibility and the emergence of non-zero magnetization. For Ising models it is proved that these two occur simultaneously. The discrepancy seems to be possible only when long-range order is impossible.

[48] This is from the graphical illustration by the editors accompanying the article "Iterated random functions" by P. Diaconis and D. Freedman, SIAM Review **41**, 45 (1999).

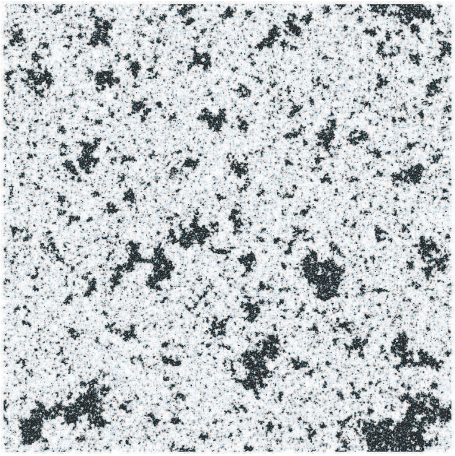


Fig. 5.6.2 A typical spin configuration of 2-Ising model exactly at $T_c$. The lattice size is 4200 × 4200 with a periodic boundary condition. If we find an even black and white composition, it is very likely that we are not *exactly* at $T_c$, because we are not observing really large fluctuations.

Even around the critical point the system does not become thermodynamically unstable. This imposes several inequalities among critical indices. For example,[49]

$$\frac{\partial(S,M)}{\partial(T,H)} \geq 0. \tag{5.6.8}$$

This inequality can be written explicitly as

$$\left.\frac{\partial S}{\partial T}\right|_H \left.\frac{\partial M}{\partial H}\right|_T \geq \left.\frac{\partial S}{\partial H}\right|_T \left.\frac{\partial M}{\partial T}\right|_H = \left.\frac{\partial M}{\partial T}\right|_H^2, \tag{5.6.9}$$

where a Maxwell's relation has been used to obtain the second equality. This implies

$$\frac{1}{T} C_H \chi \geq \left.\frac{\partial M}{\partial T}\right|_H^2. \tag{5.6.10}$$

Introducing the definitions of the critical exponents, we obtain

$$|\tau|^{-\alpha}|\tau|^{-\gamma} \geq |\tau|^{2(\beta-1)}. \tag{5.6.11}$$

Here we have ignored all the finite coefficients near the critical point (such as $T^{-1}$).[50] (5.6.11) implies that

$$|\tau|^{-(\alpha+2\beta+\gamma-2)} \geq 1 \tag{5.6.12}$$

---

[49] This requires twice differentiability of the potential, so it does not hold exactly at the critical point, but we may use it in its any neighborhood.

[50] We have assumed that the critical point is not zero; The 1-Ising model has $T_c = 0$, but this is a pathological example.

is required for $\tau \to 0$. Therefore, the quantity in the parentheses must be nonnegative:[q]

$$\alpha + 2\beta + \gamma \geq 2. \tag{5.6.13}$$

This is called *Rushbrooke's inequality*. Another example is *Griffiths's inequality* (see Problem),

$$\alpha + \beta(1 + \delta) \geq 2. \tag{5.6.14}$$

Kadanoff succeeded in elucidating the general features of critical systems with an ingenious intuitive picture.[r]

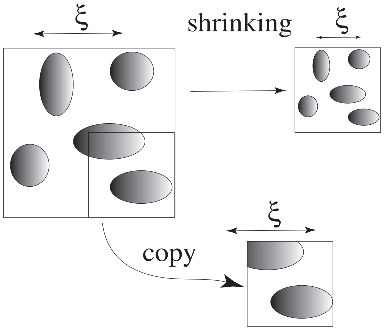


Fig. 5.6.3 The Kadanoff construction. 'Shrinking' implies looking at the system from distance with fixed eyesight, that is, coarse-graining + scaling. The outcome corresponds to the system away from the critical point, if compared with the copy of the part of the original system. $\xi$ is the correlation length.

If the original system has a temperature $\tau = (T - T_c)/T_c$ and the magnetic field $h$, then from our stepped-back point of view the system looks as if it has these parameters enlarged (away from the critical point) to $\tau\ell^{y_1}, h\ell^{y_2}$: the exponents $y_1$ and $y_2$ must be positive, where $\ell$ is the shrinking rate ($> 1$). This is a *guess* or *hypothesis*, but seems to explain everything neatly.

Let us write $m = \mathcal{M}(\tau, h)$ (this is the equation of state for the magnetic system). After one stepping-back, the volume of the region recognized as a unit cube to us would be actually the cube with edge $\ell$ (see Fig. 5.6.4) in the original system.

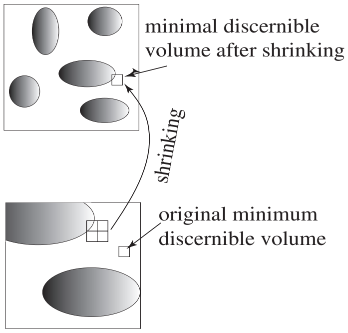


Fig. 5.6.4 If we step back and the distance between us and the sample becomes $\ell$ (in the figure $\ell = 2$) times as large as the original distance, the actual linear dimension of the minimum discernible volume becomes $\ell$-times as large as the original minimum discernible volume.

Let us put $'$ to the quantities observed after stepping back. We look at the magnetic energy stored in the minimum block $h'm'$ (after shrinking). The energy should be a much better additive quantity than the local magnetic moment, so we expect (the energy is extensive)

$$h'm' = \ell^d hm. \tag{5.6.15}$$

Since $h' = h\ell^{y_2}$, we obtain

$$m' = \mathcal{M}(\tau', h') = \ell^{d-y_2}\mathcal{M}(\tau, h) \tag{5.6.16}$$

or

$$m = \mathcal{M}(\tau, h) = \ell^{y_2-d}\mathcal{M}(\tau\ell^{y_1}, h\ell^{y_2}). \tag{5.6.17}$$

That is,**r** $\mathcal{M}$ is a generalized homogeneous function of $\tau$ and $h$.**q** This is the *scaling relation* for the equation of state. It should be clearly recognized that this is an *identity* that holds for any positive number $\ell$. Therefore, we may set $|\tau|\ell^{y_1} = 1$. Thus, we obtain from (5.6.17) ($\tau < 0$ to have non-zero magnetization)

$$m(\tau, 0) = |\tau|^{(d-y_2)/y_1} m(-1, 0). \tag{5.6.18}$$

That is,**q**

$$\beta = \frac{d - y_2}{y_1}. \tag{5.6.19}$$

We can also conclude

$$\gamma = \frac{2y_2 - d}{y_1}. \tag{5.6.20}$$

To obtain $\alpha$ we need specific heat, which is available as the second derivative of the free energy with $T$. The (singular part of the) free energy[51] $f_s = \mathcal{F}_s(\tau, h)$ per minimum discernible volume unit scales as

$$f_s = \mathcal{F}_s(\tau, h) = \ell^{-d}\mathcal{F}_s(\tau\ell^{y_1}, h\ell^{y_2}). \tag{5.6.21}$$

This comes from $f_s' = \ell^d f_s$ due to the extensivity of the free energy. If we differentiate (5.6.21) with $h$, we get (5.6.17). Historically, before Kadanoff Widom realized that if the scaling form (5.6.21) is assumed, all the critical divergences and critical index equalities can be derived. As we have seen Kadanoff furnished the key physics picture

---

[51] The free energy itself has a large nonsingular part that does not contribute to the singular behaviors near the critical point.

underlying Widom's hypothesis. Differentiating (5.6.21) twice with respect to $\tau$ (that is, $T$), we obtain

$$C(\tau, h) = \ell^{2y_1-d} C(\tau\ell^{y_1}, h\ell^{y_2}). \tag{5.6.22}$$

Therefore,

$$\alpha = \frac{2y_1 - d}{y_1}. \tag{5.6.23}$$

From (5.6.20), (5.6.19) and (5.6.23) we obtain

$$\alpha + 2\beta + \gamma = 2. \tag{5.6.24}$$

Obviously, $\xi' = \xi/\ell$ (simply shrunk), so

$$\xi(\tau.h) = \ell\xi(\tau\ell^{y_1}, h\ell^{y_2}). \tag{5.6.25}$$

Choosing $\ell = |\tau|^{-1/y_1}$, we may write this as

$$\xi = |\tau|^{-1/y_1} X_\pm(h/|\tau|^{y_2/y_1}), \tag{5.6.26}$$

where $X_\pm$ are appropriate functions for $\tau > 0$ or $\tau < 0$, so

$$\nu = 1/y_1. \tag{5.6.27}$$

This allows us to rewrite (5.6.23) as

$$\alpha = 2 - d\nu, \tag{5.6.28}$$

which is called the *hyperscaling law*.[52]

There is one more important critical exponent $\eta$ defined by the correlation function at $\tau = 0$. The correlation function for the spin $s$ is defined as

$$G(r) = \langle s(r)s(0)\rangle - \langle s(r)\rangle\langle s(0)\rangle, \tag{5.6.29}$$

where $\langle\ \rangle$ is the equilibrium average. At the critical point $\tau = h = 0$ the decay is asymptotically algebraic for large $r$:

$$G(r) \sim \frac{1}{r^{d-2+\eta}}. \tag{5.6.30}$$

---

[52] This relation (the Josephson relation) contains $d$, so we do not expect this to hold for $d > 4$. For a reason, see M. E. Fisher, Lect. Note. Phys., **186** (1993) Appendix D.

According to thermodynamic fluctuation theory,[53]

$$\chi = \beta \int d^d\boldsymbol{r} G(r) = \beta \tilde{G}(0), \tag{5.6.31}$$

where $\beta = 1/k_B T$ and $\tilde{G}$ is the Fourier transform of $G$:

$$\tilde{G}(\boldsymbol{k}) = \int d^d\boldsymbol{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} G(\boldsymbol{r}). \tag{5.6.32}$$

(5.6.31) implies that

$$\tilde{G}(0) \sim \xi^{\gamma/\nu}. \tag{5.6.33}$$

From (5.6.30) we have at $\tau = 0$ (i.e., $\xi = \infty$)

$$\tilde{G}(\boldsymbol{k}) \propto k^{\eta - 2}. \tag{5.6.34}$$

For (5.6.33) and (5.6.34) to be compatible, we must require the following functional form

$$\tilde{G}(k) \sim \xi^{\gamma/\nu} g(k\xi), \tag{5.6.35}$$

where $g$ is a well-behaved function; we have taken into account that the variable should be a dimensionless quantity $k\xi$. Furthermore, we must require that $g(x) \sim x^{-\gamma/\nu}$ to cancel $\xi$. Therefore, (5.6.34) implies

$$\frac{\gamma}{\nu} = 2 - \eta. \tag{5.6.36}$$

What happens if the system size is finite (say, its linear dimension is $L$)?**q** If $T$ is sufficiently close to $T_c$, $\xi > L$ should happen, but that is impossible; in any case, phase transition should not be clear in a finite system. Indeed, divergence is curtailed. The magnetic susceptibility reads

$$\chi = |\tau|^{-\gamma} X(h/|\tau|^{y_2/y_1}, L^{-1}/|\tau|^{1/y_1}). \tag{5.6.37}$$

For simplicity, let us suppress $h$. Even if $\tau \to 0$, the correlation length cannot become larger than the system size; $\xi \to \infty$ is not realizable. Therefore, $\chi$ should not depend

---

[53] $\beta\langle M^2\rangle/V$ is the susceptibility per volume, which we can interpret as $\chi$. Now,

$$\langle M^2\rangle = \left\langle \left(\int d^d\boldsymbol{r} S(r)\right)^2 \right\rangle = V \int d^d\boldsymbol{r} G(r).$$

on $\xi$ (or on $\tau$). Then, $X(0,x)$ must be a power function of $x$ in the $x \to \infty$ limit (the standard scaling argument). The power is determined to cancel $\tau$ in (5.6.37):

$$\chi \sim (L^{-1})^{-\gamma/\nu} = L^{\gamma/\nu} = L^{2-\eta}, \tag{5.6.38}$$

where we have used $\nu = 1/y_1$. The lack of divergence is consistent with the curtailed fluctuation. This corresponds to the roundoff of the phase transition as expected.[54]

## 5.7 Renormalization group flow

Kadanoff's idea (Kadanoff construction) consists of two parts: coarse-graining and scaling. The crux of the idea is: if the system is at the critical point, then the configuration is invariant under coarse-graining $\mathcal{K}$ and scaling $\mathcal{S}$. That is, if we define $\mathcal{R} = \mathcal{K}\mathcal{S}$, then thermodynamic observables are invariant (up to the change due to extensivity or intensivity of the observables) under the application of $\mathcal{R}$. To apply $\mathcal{R}$ is to observe the system from distance with a fixed eyesight.[55] Fig. 5.7.1 illustrates how iterative operation of $\mathcal{R}$ drives the statistical configurations at various temperatures.

Operating $\mathcal{R}$ is called a *renormalization group transformation*. We can understand its iterative applications as product of $\mathcal{R}$, and the unit element corresponds to doing nothing to the system, so the totality of the renormalization transformations makes a monoid, but informally we call it a *renormalization group*. According to Kadanoff's original idea, the image due to $\mathcal{R}$ is the same system under a different condition (e.g., at a different temperature), so we may understand that $\mathcal{R}$ transforms a thermodynamic state into another (of the same material system); we may imagine that successive applications of $\mathcal{R}$ defines a flow (*renormalization group flow* = RG

---

[54] M-C Wu and C-K Hu, "Exact partition functions of the Ising model on $M \times N$ planar lattices with periodic-aperiodic boundary conditions," J. Phys., A **35**, 5189-5206 (2002) gives exact solutions to the 2D Ising model on finite lattice with various boundary conditions. We can find explicit results for the roundoff.

[55] ⟪**History and renormalization**⟫ "It is true that it is only possible to anticipate the general structure of the future, but that is all that we in truth understand of the past or of the present. Accordingly, if you want a good view of your own age, look at it from far off. From what distance? The answer is simple. Just far enough to prevent you seeing Cleopatra's nose." (Ortega, *La rebelión de las masas* (1930) p55)

flow) on the phase diagram.[56]**r**

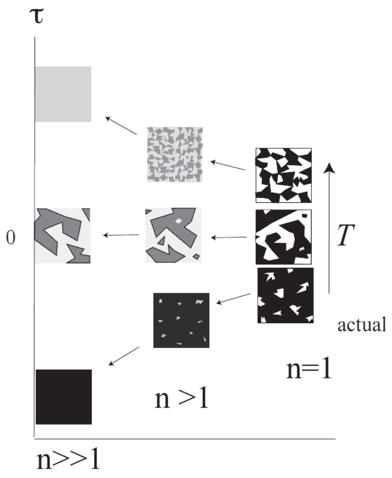


Fig. 5.7.1 The result of Kadanoff construction, or the real space renormalization group transformation. For simplicity, $h = 0$. Here, $\tau = (T - T_C)/T_C$ and $n$ is the number of operations of renormalization group transformation $\mathcal{R}$. The vertical line on the left is the destination after many applications of $\mathcal{R}$. Only when the starting point is right, the system can stay within the finite range of $\tau \sim 0$. The low temperature state is driven to one of the ordered phases at $T = 0$; this time, it happens to be totally 'up.' If the starting point is $T > T_c$, the state is driven to $T = \infty$ state.

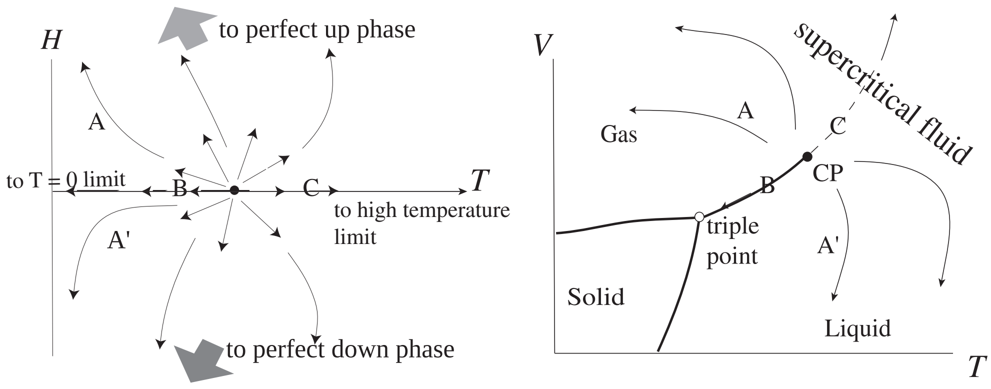


Fig. 5.7.2 RG flows on the phase diagram. Left: (Ising) magnet. There are five ultimate destinations (high temperature limit, phase boundary, critical point, all up and all down low temperature states); Right: Fluid system (actually, the flow corresponding to that of the magnetic system is defined for fluid phases only.)

At the fixed point $\mathcal{R}\xi = \xi$. Since $\mathcal{S}$ definitely shrinks the system, this condition is satisfied only if $\xi = 0$ or $\xi = \infty$. That is, the phases without spatial correlation at all or critical points are the only possible fixed points. Notice that if we understand

---

[56] As we will see soon, the RG flow does not generally flow on the phase diagram (of a given material). In terms of Fig. 5.7.1, the phase diagram corresponds to what happens on the $n = 0$ slice. The actual flows move as $n = 1, 2, \cdots$. The flows in Fig. 5.7.2 are, intuitively, the projection of these RG flow lines onto $n = 1$.

these fixed points, we understand the general structure of the phase diagram. The ordinary bulk phases from our macroscopic point of view do not have any appreciable correlation distance, so $\xi = 0$ fixed points are bulk phases. To understand their macroscopic properties we need not worry (qualitatively) about spatial correlations of fluctuations. This is the reason why the so-called mean-field theory is useful, which will be discussed shortly (see Section 5.8). Thus, to understand the phase diagram, we use mean field theory to understand the bulk phases not too close to the critical points,[57] and use renormalization group theory to understand the features near the critical points.

We may interpret the renormalization group transformation as a map from a (generalized) canonical distribution $\mu$ to another (generalized) canonical distribution $\mu' = \mathcal{R}\mu$. We can imagine an effective Hamiltonian $H$ (it is customary that $\beta$ is absorbed in $H$) such as

$$\mu = \frac{1}{Z}e^{-H},\ \mu' = \frac{1}{Z'}e^{-H'}. \tag{5.7.1}$$

We may write $H' = \mathcal{R}H$. Therefore, we can imagine that successive applications of $\mathcal{R}$ defines a flow in the space of Hamiltonians (or models or systems). This idea is illustrated in Fig. 5.7.3.

In Fig. 5.7.3 $H^*$ is a fixed point with an infinite correlation length of the RG flow. Its stable manifold $W_S(H^*)$ is called the *critical surface*. The Hamiltonian of the actual material, say, magnet A, changes as the temperature changes along the trajectory denoted by the curve with 'magnet A.' It crosses the critical surface at its critical temperature. The renormalization transformation uses the actual microscopic Hamiltonian of magnet A at various temperatures as its initial conditions. Three representative RG flows for magnet A are depicted. 'a' is slightly above the critical temperature, 'b' exactly at $T_c$ of magnet A ('b′' is the corresponding RG trajectory for magnet B), 'c' slightly below the critical temperature. Do not confuse the trajectory of the actual microscopic system as temperature changes and the trajectories produced by the RG transformation.

> **Some technical terms of dynamical systems**[58]
> The theory of dynamical systems studies the flow induced by the map on a certain spaces, so the technical terms of this theory is convenient to describe renormalization

---

[57] This does not mean that we can use the original microscopic Hamiltonian when we utilize a mean-field approach; we must use an appropriately renormalized Hamiltonian. Therefore, a precise statement is: there is a Hamiltonian (with short-range interactions) that can be used to describe the macroscopic features of a bulk phase with the aid of a mean-field approach. See the following explanation of the Hamiltonian flow.

[58] J. Palis, Jr. and W. de Melo, *Geometric Theory of Dynamical Systems, an introduction* (Springer, 1982) is a good introductory book.

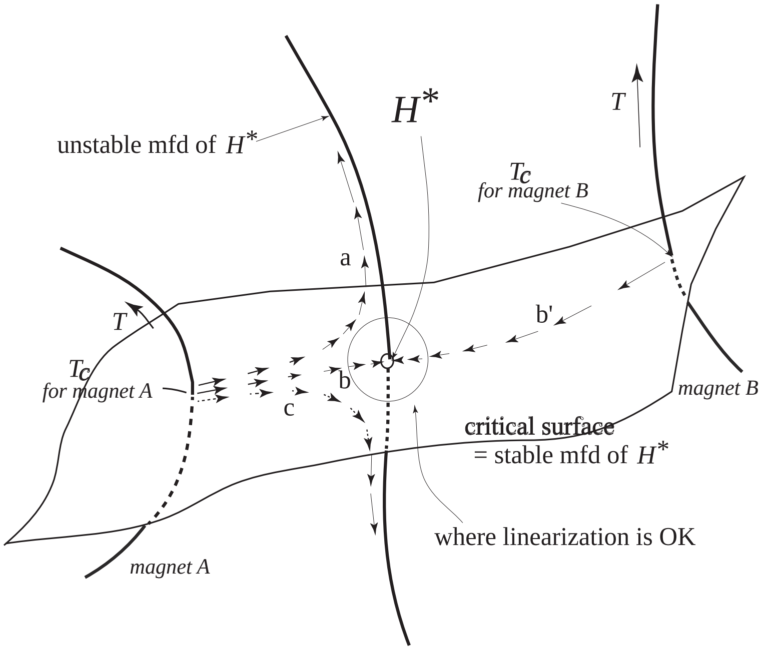


Fig. 5.7.3**r** A global picture**r** of renormalization group flow in the Hamiltonian space $\mathcal{H}$. The explanation is in the text. ‘mfd’ = manifold.

group flows. If $T$ is a map from a set $A$ into itself, $(T, A)$ is a (discrete time) *dynamical system*. A point $x \in A$ such that $Tx = x$ is called a *fixed point* of the dynamical system. There may not be any such point or there may be infinitely many of them. For a fixed point $x$, the set $W_s(x)$ of points $y \in A$ such that $\lim_{n\to\infty} T^n y = x$ is called the *stable manifold* of $x$. The linear approximation of $T$ around a fixed point $x$ is determined by the derivative $D_xT$ at $x$. The eigenspace of eigenvalues of $D_xT$ inside the unit disk is tangent to the stable manifold of $x$ at $x$. The set $W_u(x)$ of points $y$ such that $T^{-n}y \to x$ is called the *unstable manifold* of $x$.[59] This is tangent to the eigenspace of the eigenvalues of $D_xT$ outside the unit disk. If the dimension of the stable manifold is zero, the fixed point is called a *source*; if the dimension of the unstable manifold is zero, the fixed point is called a *sink*.

If $D_xT$ has no eigenvalue on the unit circle, the fixed point $x$ is called a *hyperbolic fixed point*. Around a hyperbolic fixed point, the flow induced by the map $T$ may be approximated by the flow induced by its linearization $D_xT$. More precisely, there is a homeomorphism (= one to one map continuous in both directions) $\phi$ such that $\phi(T(y)) = D_xT(\phi(y))$ for $y$ in an appropriate neighborhood of $x$ (the *Hartman-Grobman theorem*).

Figure 5.7.3 tells us that to understand the critical phenomenon we have only to

---

[59] Here, for simplicity, $T^{-1}$ is also considered as a map; if $T^{-1}$ is not a map $T^{-n}y \to x$ should read $x \in T^{-n}y$.

study $H^*$ and its neighborhood. The fixed point is usually hyperbolic (see the fine-lettered explanation above), so we can study the behavior of RG near $H^*$ through linear approximation. We will come back to this problem after mean field theory.

According to the Kadanoff construction, after coarse-graining the 'spin' in the coarse-grained Hamiltonian must be, as can be guessed from the illustrations, some sort of 'blocked spins' (called *block spins*). Now, let us go back to basic statistics of iid stochastic variables $\{X_i\}$. The pillars of modern probability theory are the law of large numbers and its refinements. Statistical mechanics heavily relies on the law of large numbers. The thermodynamic theory of fluctuation is an application of large deviation theory. The other major refinement of the law of large numbers is the central limit theorem, which has not really been used yet in this book.[60] The central limit theorem for the iid stochastic variables with zero average reads that the distribution function of $(1/\sqrt{N})\sum_{i=1}^{N} X_i$ converges to $N(0,V)$, where $V$ is the variance of $X_1$. The partial sum $S_N = \sum_{i=1}^{N} X_i$ corresponds to the block spin in Kadanoff's picture, so making the partial sum is a sort of coarse-graining. Dividing with $\sqrt{N}$ corresponds to scaling. The fixed point distribution of the 'renormalization group transformation' $\mathcal{R} : X_1 \to S_N/\sqrt{N} = (1/\sqrt{N})\sum_{i=1}^{N} X_i$ is $N(0,V)$. Here, the scaling factor $\sqrt{N}$ is chosen carefully to have a nontrivial limit: not $\delta$-function-like nor zero everywhere.[61]

Thus, we see renormalization group theory is the central limit theorem$_{\mathbf{r}}$ for strongly correlated stochastic variables.[62]

---

[60] **⟪Central limit theorem vs. large deviation⟫** The reader might claim that it has already been used to understand fluctuations; isn't the Gaussian nature of fluctuation the sign of central limit theorem? This is only accidental for short-correlated systems. Fluctuation studies the deviation of the average from the true average, when the system size is small. We ask how the fluctuation *of* the mean *disappears* as the system size increases. In contrast, the central limit theorem is concerned with small deviations *from* the mean that appropriately scales with the system size.

[61] The authoritative reference of the central limit theorem for iid variables is W. Feller, *An introduction to probability theory and its applications*, vol. II (Wiley, 1971). The ideas of flow, fixed point, stability, etc., all appeared in the classical theory long before renormalization group theory.

[62] This view point was clearly recognized first by G. Jona-Lasinio, "The renormalization group: a probabilistic view," Nuovo Cim. **268**, 99-119 (1975).

## 5.8 Mean field theory

As discussed in the preceding section to understand the phase diagram globally, we may ignore the correlation effects except near critical points, where the correlation length diverges. Away from critical points, if we wish to compute the equilibrium average of a function of several spins $f(s_0, s_1, \cdots, s_n)$ we may separately average all the spins. Furthermore, if we assume $\langle s_i^k \rangle \sim \langle s_k \rangle^k$ (i.e., if we assume that fluctuations are not large), we arrive at

$$\langle f(s_0, s_1, \cdots, s_n) \rangle \simeq f(\langle s_0 \rangle, \langle s_1 \rangle, \cdots, \langle s_n \rangle). \tag{5.8.1}$$

This is the fundamental idea of the *mean field* approach. Here, let us proceed slightly more systematically.

Let us recall an elementary identity of probability theory.**r** If $\cup_i B_i = \Omega$ and $B_i \cap B_j = \emptyset$ for $i \neq j$ (i.e., $\{B_i\}$ is a partition of the total event), then

$$E\left(E(A|B_i)\right) = E(A); \tag{5.8.2}$$

That is, the average of a conditional expectations over all the conditions is equal to the unconditional average.

Let us choose as $B$ a particular configuration $\{s_1, \cdots, s_{2d}\}$ of all the spins interacting with the 'central spin' $s_0$ on a $d$-cubic lattice.

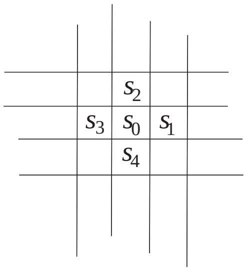


Fig. 5.8.1 The central spin $s_0$ and its nearest neighbor surrounding spins $s_1, \cdots, s_{2d}$.

We can compute the following conditional average for the $d$-Ising model on the (hyper)cubic lattice exactly:

$$E(s_0|s_1, \cdots, s_{2d}) = \frac{\sum_{s_0} s_0 e^{\beta J s_0(s_1+\cdots+s_{2d})+\beta h s_0}}{\sum_{s_0} e^{\beta J s_0(s_1+\cdots+s_{2d})+\beta h s_0}} = \tanh[\beta h + \beta J(s_1 + \cdots + s_{2d})]. \tag{5.8.3}$$

Because $E(s_0) = E\left(E(s_0|s_1, \cdots s_{2d})\right)$, we obtain**q**

$$\langle s_0 \rangle = \langle \tanh\left[\beta h + \beta J(s_1 + \cdots + s_{2d})\right] \rangle. \tag{5.8.4}$$

This is an *exact* relation into which we may introduce various approximations to construct mean field approaches.

Now, to compute the RHS of (5.8.4), we must introduce some approximation. The most popular (and simple-minded) version is (5.8.1):

$$\langle \tanh[\beta h + \beta J(s_1 + \cdots + s_{2d})] \rangle \simeq \tanh[\beta h + \beta J \langle s_1 + \cdots + s_{2d} \rangle]. \tag{5.8.5}$$

Therefore, for $m = \langle s_0 \rangle$, we obtain a closed equation

$$m = \tanh[\beta(2dJm + h)]. \tag{5.8.6}$$

$2dJm$ may be understood as an effective field acting on $s_0$, so this is called the *mean field* (sometimes called the *molecular field* as well). This is the etymology of the name of the approximation method being considered. Let $2d\beta Jm = x$. (5.8.6) reads

$$x = 2d\beta J \tanh(x + \beta h). \tag{5.8.7}$$

For simplicity, let us assume $h = 0$. We have to solve

$$x = 2d\beta J \tanh x. \tag{5.8.8}$$

This may be graphically solved (Fig. 5.8.2).

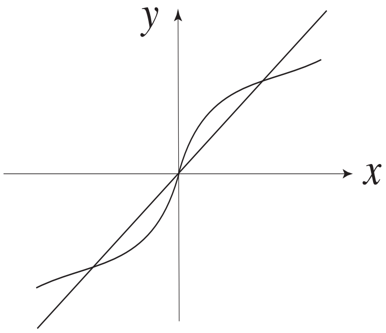


Fig. 5.8.2 The solution to (5.8.8) may be obtained graphically.

The bifurcation[63] from the case with a single solution to that with 3 solutions occurs at $2d\beta J = 1$. That is, this gives the phase transition temperature $T_c$. $m$ increases as $|T - T_c|^{1/2}$ (i.e., the critical exponent $\beta = 1/2$).

---

[63] A phenomenon that the solution changes its character is called *bifurcation*. There are many types, and this is a *pitchfork bifurcation*; if we know this, the exchange of the stability of the branches immediately tells us the stabilities of the branches as illustrated in the text.

To conclude that the bifurcation actually signifies the phase transition (within the mean-field approximation), we must check that the nonzero solutions are the equilibrium solutions. That is, we must demonstrate that the $m \neq 0$ solution has a lower free energy than the $m = 0$ case. The best way may be to study the bifurcation diagram and check the stability of the solution under small perturbations. The stability of $m \neq 0$ state is obvious.

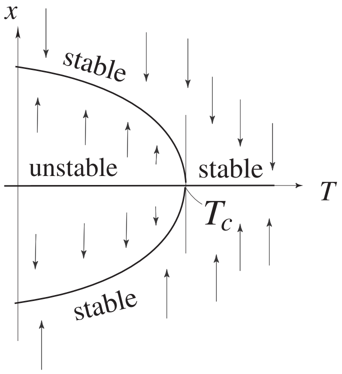


Fig. 5.8.3 The stability of the solution to (5.8.8) may also be understood graphically.

To obtain the equation of state, we must include the external magnetic field $h$. From (5.8.6) we can obtain

$$\beta h = \operatorname{Arctanh} m - 2d\beta J m = (1 - 2d\beta J)m + \frac{1}{3}m^3 + O[m^3]. \tag{5.8.9}$$

Therefore, we obtain $\delta = 3$, $\gamma = 1$ and $\beta = 1/2$. Combining these results and assuming Rushbrooke's equality (5.6.13), we conclude $\alpha = 0$. As we have already discussed, the mean field theory is not designed to describe critical phenomenon, so these critical exponents need not be very realistic.

A better approximation to (5.8.4) could be invented. We know $s = \pm 1$. Therefore, $s_i^2 = 1$ (no summation convention), so, if we write the parity of $x$ as $\sigma(x)$ (0, if $x$ is even and 1, if $x$ is odd), $s_1^p s_2^q s_3^r s_4^s = s_1^{\sigma(p)} s_2^{\sigma(q)} s_3^{\sigma(r)} s_4^{\sigma(s)}$. Since $\tanh x$ is an odd analytic function, we may write as ($2d$-case, for simplicity)

$$\tanh \beta J(s_1 + s_2 + s_3 + s_4) = a(s_1 + s_2 + s_3 + s_4) + b(s_1 s_2 s_3 + s_2 s_3 s_4 + s_3 s_4 s_1 + s_4 s_1 s_2), \tag{5.8.10}$$

where $a$ and $b$ are numerical coefficients. This relation must be an identity. Hence, $a$ and $b$ may be chosen by computing particular spin configurations. All up gives us

$$\tanh 4\beta J = 4a + 4b, \tag{5.8.11}$$

while only one down gives us

$$\tanh 2\beta J = 2a - 2b. \tag{5.8.12}$$

All other configurations give one of the above, or $0 = 0$. We get $a = (\tanh 4\beta J + 2\tanh 2\beta J)/8$ and $b = (\tanh 4\beta J - 2\tanh 2\beta J)/8$. Writing $\langle s \rangle = m$ and assuming the spatial uniformity, (5.8.4) reads

$$m = 4am + 4b\langle s_1 s_2 s_3 \rangle. \tag{5.8.13}$$

This is still exact. Therefore, we should realize that for 1D this method gives an exact equation for the magnetization.

If we approximate $\langle s_1 s_2 s_3 \rangle = m^3$, we have a closed equation

$$m = 4am + 4bm^3. \tag{5.8.14}$$

This gives us a considerable improvement over simple mean field theory.

**Warning**. We have introduced the idea of mean field theory to study the system thermodynamic sufficiently away from critical points. Therefore, the mean field theory cannot generally assert anything about the phase transition. It cannot guarantee the existence of phase transition (esp., second order phase transition) even if it concludes that there is one. Recall that even for $d = 1$, the mean field theory (a simple version) asserts that there is a second order phase transition at some finite $T$. We know this cannot be true. Even in the case where a phase transition occurs, it cannot reliably predict whether the phase transition is continuous or not. However, if fluctuation effects are not serious, then the mean field results become qualitatively reliable. Thus, it is believed that if $d \geq 4$ (especially $d > 4$), for fluids and magnets, the simplest mean field results are generally qualitatively correct.[64,65]

However, if a mean field theory concludes that there is no ordering phase transition, this conclusion sounds very plausible. Since mean field theory ignores fluctuations, it should overestimate the ordering tendency. For the ferromagnetic Ising

---

[64] ⟪**Mean field results are reliable in high dimensional spaces**⟫ Recent progress in this respect is: M. Biskup and L. Chase, "Rigorous analysis of discontinuous phase transitions via mean-field bounds," Commun. Math. Phys. **238**, 53 (2003). Under a technical condition (the reflection positivity) it is shown that whenever the mean-field theory predicts a discontinuous transition, the actual model also undergoes a discontinuous transition (which occurs near the mean-field transition temperature), provided the spatial dimension is sufficiently high and/or the phase transition is sufficiently strong.

[65] ⟪**Strong first order phase transtions may be predicted by mean field**⟫ L. Chayes, "Mean Field Analysis of LowDimensional Systems," Commun. Math. Phys., **292**, 303 (2009) tells us that if the order parameter jumps sufficiently at the transition, the mean field theory can predict the first order phase transition correctly even in 2 and 3-spaces.

model this expectation has been vindicated.[66] The same idea tells us that the mean field critical temperature should be the upper bound of the true critical temperature: $T_c \le T_{c,mean}$.

## 5.9 Transfer matrix

**q**

Before going to an actual renormalization group illustration, let us look at a technique called the *transfer matrix method*. This was the method used by Onsager to evaluate the partition function of the 2-Ising model on the square lattice.[67]

Let us consider a 1-Ising model with the Hamiltonian (+ the interaction Hamiltonian with the magnetic field $h$) given by

$$H = -J \sum_{i=1}^{N-1} s_i s_{i+1} - h \sum_{i=1}^{N} s_i. \tag{5.9.1}$$

Let us define the partition function $Z_N(+)$ for the length $N$ spin chain with the $N$th spin up:

$$Z_N(+) = \sum_{\{s_n\}_{n=1}^{N-1}} e^{\beta[J(+1)s_{N-1}+h(+1)]} e^{\beta[Js_{N-1}s_{N-2}+hs_{N-1}]} \cdots e^{\beta[Js_2 s_1 + hs_2]} e^{\beta h s_1}. \tag{5.9.2}$$

We can analogously define the partition function $Z_N(-)$ for the length $N$ spin chain

---

[66] C. J. Thompson, "Upper bounds for Ising model correlation functions," Commun. math. Phys., **24**, 61 (1971) proves that, if the magnetic field is non-negative, then the mean field magnetization density is the upper bound of the true magnetization density.

[67] ⟪**Onsager's biography**⟫ See C. Longuet-Higgins and M. E. Fisher, "Lars Onsager: November 27, 1903-October 5, 1976," J. Stat. Phys., **78**, 605 (1995). This is Onsager's biography everyone can enjoy. According to this, Onsager applied the transfer matrix method to the strip of width 2, 3 and 4 lattice points, and constructed a conjecture from these results, then confirmed it for the width 5 strip and closed in on the general formula. "His statistical mechanics were popularly known as 'Advanced Norwegian I' and 'Advanced Norwegian II'." He was fired more than once for his poor teaching, and his Nobel-prize winning dissertation intended for his PhD was rejected as insufficient from his alma mater.

with the $N$th spin down. In terms of these, we can make $Z_{N+1}(+)$ as

$$Z_{N+1}(+) = \sum_{s=\pm 1} e^{\beta[Js+hs]} Z_N(s). \tag{5.9.3}$$

Therefore, if we introduce the vector

$$\boldsymbol{Z}_N = \begin{pmatrix} Z_N(+) \\ Z_N(-) \end{pmatrix}, \tag{5.9.4}$$

$$\boldsymbol{Z}_{N+1} = \boldsymbol{T}\boldsymbol{Z}_N, \tag{5.9.5}$$

where $\boldsymbol{T}$, called the *transfer matrix*,[68]**r** is defined as

$$\boldsymbol{T} = Matr(e^{\beta[Jss'+hs]}) = \begin{pmatrix} e^{\beta J+\beta h} & e^{-\beta J+\beta h} \\ e^{-\beta J-\beta h} & e^{\beta J-\beta h} \end{pmatrix}. \tag{5.9.6}$$

Notice that

$$Z_N = (1,1)\boldsymbol{Z}_N. \tag{5.9.7}$$

Repeated use of the recursion (5.9.5) results in

$$\boldsymbol{Z}_N = \boldsymbol{T}^N \begin{pmatrix} e^{\beta h} \\ e^{-\beta h} \end{pmatrix}. \tag{5.9.8}$$

In this case the first spin is free to point up or down. For a ring of $N$ spins ($s_1 = s_{N+1}$), as we see immediately, $Z_N = Tr\boldsymbol{T}^N$.

The easiest method to compute (5.9.8) is to use a similarity transformation (or unitary transformation, if possible) to convert $\boldsymbol{T}$ into a diagonal form:[69]

$$\boldsymbol{T} = U^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U, \tag{5.9.9}$$

where $\lambda_1$ and $\lambda_2$ are eigenvalues of $\boldsymbol{T}$, and $U$ is the orthogonal transformation needed to diagonalize $\boldsymbol{T}$. Introducing (5.9.9) into (5.9.8), we obtain

$$\boldsymbol{Z}_N = U^{-1} \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} U \begin{pmatrix} e^{\beta h} \\ e^{-\beta h} \end{pmatrix}. \tag{5.9.10}$$

---

[68] ⟪**The origin of the transfer matrix method**⟫ The method was devised by Kramers and Wannier: Phys. Rev. **60**, 252 (1941). For a continuum model, an integral equation approach can be used and was devised by H. Takahashi almost simultaneously in 1942 (Proc. Phys-Math. Soc. Japan **24**, 60 (1942)). He showed that 1D short-range systems cannot have any phase transition for $T > 0$. [In 1941 Japan attacked Pearl Harbor; in 1942 Fermi and collaborators succeeded in nuclear chain reaction.]

[69] if impossible, in a Jordan normal form; in the present case, the eigenvalues are distinct, so the matrix is diagonalizable.

Therefore, we finally have the following structure:

$$Z_N = a\lambda_1^N + b\lambda_2^N, \tag{5.9.11}$$

where $a$ and $b$ are real numbers. If we assume $\lambda_1 > |\lambda_2|$, $a$ is positive, and, since $N \gg 1$, the first term dominates $Z_N$. Therefore, the free energy per spin is given by

$$f = -k_B T \log \lambda_1. \tag{5.9.12}$$

Depending on the boundary conditions, the exact formula for the partition function changes, but the free energy per spin (this is the only quantity meaningful in the thermodynamic limit $N \to \infty$) depends only on the largest eigenvalue of the transfer matrix that is not dependent on the boundary condition.

The elements of the transfer matrix are entire functions[70] of $T$ or $h$, and eigenvalues are their *algebraic functions*.[71] Therefore, as long as eigenvalues are finite, their singularities are branch points. The branch points of the eigenvalues occur when they change their multiplicities (digeneracies), so the multiplicity of the largest eigenvalue is of vital importance. The key theorem we need is the following famous and important theorem:

**Theorem [Perron and Frobenius]$_{\mathbf{q}}$**
Let $A$ be a square matrix whose elements are all non-negative, and there is a positive integer $n$ such that all the elements of $A^n$ are positive. Then, there is a non-degenerate real positive eigenvalue $\lambda$ such that
(i) $|\lambda_i| < \lambda$, where $\lambda_i$ are eigenvalues of $A$ other than $\lambda$,[72]
(ii) the elements of the eigenvector belonging to $\lambda$ may be chosen all positive. □
This special real eigenvalue giving the spectral radius is called the *Perron-Frobenius eigenvalue.*

**Proof of the Perron-Frobenius theorem**[73]
Let us introduce the vectorial inequality notation: $\boldsymbol{x} > 0$ ($\geq 0$) implies that all the components of $\boldsymbol{x}$ are positive (non-negative). Also let us write $\boldsymbol{x} \geq (>) \boldsymbol{y}$ if $\boldsymbol{x} - \boldsymbol{y} \geq (>) 0$.
Let $\boldsymbol{x}$ be a vector such that $|\boldsymbol{x}| = 1$ and $\boldsymbol{x} \geq 0$. The largest $\rho$ satisfying

$$A\boldsymbol{x} \geq \rho \boldsymbol{x} \tag{5.9.13}$$

---

[70] A function that is holomorphic except at infinity is called an *entire function.*

[71] An algebraic function of $\{x_i\}$ is a function constructed from $\{x_i\} \cup \boldsymbol{R}$ by applying finitely many times the elementary operations (= addition, multiplication, division, subtraction, and integer root extraction) on these elements.

[72] That is, $\lambda$ gives the spectral radius of $A$.

[73] A standard reference may be E. Seneta, *Non-negative matrices and Markov chains* (Springer, 1980). The proof here is an eclectic version due to many sources, including N. Iwahori, *Graphs and Stochastic Matrices* (Sangyo-tosho, 1974).

is denoted by $\Lambda(\boldsymbol{x})$. The proof is divided into several steps.
(i) Since the set $U = \{\boldsymbol{x} \,|\, \boldsymbol{x} \geq 0, |\boldsymbol{x}| = 1\}$ is a compact set,[74] there is a vector $\boldsymbol{z} \in U$ that maximizes $\Lambda(\boldsymbol{x})$. Let us write $\lambda = \Lambda(\boldsymbol{z})$.
(ii) $\lambda$ is an eigenvalue of $A$, and $\boldsymbol{z}$ belongs to its eigenspace: $A\boldsymbol{z} = \lambda\boldsymbol{z}$.
[Demo] Even if not, we have $\boldsymbol{w} = A\boldsymbol{z} - \lambda\boldsymbol{z} \geq 0$ (not equal to zero). Notice that for any vector $\boldsymbol{x} \geq 0$ but $\neq 0$ there is some positive integer $m$ such that $A^m\boldsymbol{x} > 0$, so

$$A^m\boldsymbol{w} = AA^m\boldsymbol{z} - \lambda A^m\boldsymbol{z} > 0. \tag{5.9.14}$$

This implies $\Lambda(A^m\boldsymbol{z}) > \lambda$, but $\lambda$ is the maximum of $\Lambda$, this is a contradiction. Therefore, $\boldsymbol{w} = 0$. That is, $\boldsymbol{z}$ is an eigenvector belonging to $\lambda$.
(iii) We may choose $\boldsymbol{z} > 0$.
[Demo] $z \geq 0$ and nonzero, so there is a positive integer $m$ such that $A^m\boldsymbol{z} > 0$ but this is $\lambda^m\boldsymbol{z} > 0$, so actually $\boldsymbol{z} > 0$.
(iv) $\lambda$ is the spectral radius of $A$.
[Demo] Suppose $A\boldsymbol{y} = \lambda'\boldsymbol{y}$. Let $\boldsymbol{q}$ be the vector whose components are absolute values of $\boldsymbol{y}$: $q_i = |y_i|$. Then, $A\boldsymbol{q} \geq |\lambda'|\boldsymbol{q}$. Therefore, $|\lambda'| \leq \lambda$.
(v) The absolute value of other eigenvalues are smaller than $\lambda$. That is, no eigenvalues other than $\lambda$ is on the spectral circle.
[Demo] Suppose $\lambda'$ is an eigenvalue on the spectral circle but is not real positive. Let $\boldsymbol{q}$ be the vector whose components are absolute values of an eigenvector belonging to $\lambda'$. Since $A\boldsymbol{q} \geq |\lambda'|\boldsymbol{q} = \lambda\boldsymbol{q}$, actually we must have $A\boldsymbol{q} = \lambda\boldsymbol{q}$. That is, the absolute value of each component of the vector $A^m\boldsymbol{y} = \lambda'^m\boldsymbol{y}$ coincides with the corresponding component of $A^m\boldsymbol{q}$. This implies

$$\left|\sum_j (A^m)_{ij} y_j\right| = \sum_j (A^m)_{ij}|y_j| = \sum_j |(A^m)_{ij} y_j|. \tag{5.9.15}$$

All the components of $A^m$ are real positive, so all the arguments of $y_j$ are identical.[75] Hence, $\lambda' = \lambda$.
(vi) $\lambda$ is non-degenerate.
[Demo] Suppose otherwise. Then, there is a vector $\boldsymbol{z}'$ that is not proportional to $\boldsymbol{z}$ but still $A\boldsymbol{z}' = \lambda\boldsymbol{z}'$. Here, $A$ is a real matrix and $\lambda$ is real, we may choose $\boldsymbol{z}'$ to be real. Since $\boldsymbol{z}$ and $\boldsymbol{z}'$ are not parallel, we may choose $\alpha$ appropriately so that $\boldsymbol{v} = \boldsymbol{z} + \alpha\boldsymbol{z}' \geq 0$ but has a zero component. This is contradictory to (iii). □

Since the transfer matrix is with positive elements, the logarithm of its Perron-Frobenius eigenvalue gives the free energy per spin. If the number of states for each 1D element is finite and the interaction range is finite, then no phase transition occurs for $T > 0$, because the transfer matrix is finite dimensional.

What happens if $T \to 0$ for the 1-Ising model?**r** Let us explicitly compute the

---

[74] The domain of $\lambda$ is a sphere of some dimension restricted to the nonnegative coordinate sector. Thus, the problem is to find the largest value of a continuous function $\lambda$ on a closed set. The maximum value theorem guarantees the existence of the maximum point.

[75] $a, b \neq 0$ and $|a + b| = |a| + |b|$ imply the real positivity of $a/b$.

eigenvalues of the transfer matrix (5.9.6):

$$\lambda_\pm = e^{\beta J}\left[\cosh\beta h \pm \sqrt{\sinh^2\beta h + e^{-4\beta J}}\right]. \tag{5.9.16}$$

Unless $e^{-4\beta J}$ vanishes, they are analytic in $h$, so no spontaneous magnetization is possible. That is, there is no spontaneous magnetization for $T > 0$.

The ratio $\lambda_-/\lambda_+$ reads without magnetic field ($h = 0$)

$$\frac{\lambda_-}{\lambda_+} = \frac{1 - e^{-2\beta J}}{1 + e^{-2\beta J}}. \tag{5.9.17}$$

Therefore, in the $T \to 0$ limit, $\beta J \to \infty$ and the ratio converges to unity. Its implication can be understood from the correlation function (notice that $\langle s\rangle = 0$)

$$\langle s_{i+r}s_i\rangle \sim (\lambda_-/\lambda_+)^r. \tag{5.9.18}$$

That is, the correlation length diverges in the $T \to 0$ limit. That is, for 1-Ising model $T_c = 0$.

**Correlation and eigenvalues of transfer matrix**

The correlation function $\langle s_i s_j\rangle$ (without magnetic field) may be calculated as follows:

$$\langle s_i s_j\rangle = \frac{1}{Z_N}\sum_{\{s_n\}_{n=1}^{N+1}} e^{\beta J s_{N+1}s_N}e^{\beta J s_N s_{N-1}}\cdots e^{\beta J s_{i+1}s_i}s_i e^{\beta J s_i s_{i-1}}\cdots e^{\beta J s_{j+1}s_j}s_j e^{\beta J s_j s_{j-1}}\cdots e^{\beta J s_2 s_1}. \tag{5.9.19}$$

This formula can be rewritten as

$$\langle s_{i+r}s_i\rangle = \frac{1}{Z_N}(1,1)\boldsymbol{T}^{N-i-r}\sigma\boldsymbol{T}^{r+1}\sigma\boldsymbol{T}^{i-1}\begin{pmatrix}1\\1\end{pmatrix}, \tag{5.9.20}$$

where

$$\sigma = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{5.9.21}$$

Using the diagonalization of $\boldsymbol{T}$, we have

$$\langle s_{i+r}s_i\rangle = \frac{1}{Z_N}(1,1)[U^{-1}\Lambda^{N-i-r}U]\,\sigma\,[U^{-1}\Lambda^{r+1}U]\,\sigma\,[U^{-1}\Lambda^{i-1}U]\begin{pmatrix}1\\1\end{pmatrix}. \tag{5.9.22}$$

To compute this further, we note that the magnetization is zero. That is, for any $j$

$$\langle s_j\rangle = \frac{1}{Z_N}(1,1)U^{-1}\left[\Lambda^{N-j}U\sigma U^{-1}\Lambda^j\right]U\begin{pmatrix}1\\1\end{pmatrix} = 0 \tag{5.9.23}$$

Now, pay attention to the factor in the square brackets:

$$\left(\begin{array}{cc} \lambda_+^{N-j} & 0 \\ 0 & \lambda_-^{N-j} \end{array}\right) U\sigma U^{-1} \left(\begin{array}{cc} \lambda_+^{j} & 0 \\ 0 & \lambda_-^{j} \end{array}\right) \tag{5.9.24}$$

This is dominated by

$$\lambda_+^{N-j}[U\sigma U^{-1}]_{11}\lambda_+^j \tag{5.9.25}$$

so for any $i$ (recall that $Z_N \sim \lambda_+^N$)

$$\langle s_i \rangle \simeq \frac{1}{Z_N}\lambda_+^N[U\sigma U^{-1}]_{11} \simeq [U\sigma U^{-1}]_{11} = 0. \tag{5.9.26}$$

This implies that

$$\langle s_{i+r}s_i \rangle \simeq \frac{1}{Z_N}(1,1)U^{-1}\left\{\lambda_+^{N-i-r}[U\sigma U^{-1}]_{12}\lambda_-^r[U\sigma U^{-1}]_{21}\lambda_+^{i-1}\right\}U\left(\begin{array}{c} 1 \\ 1 \end{array}\right). \tag{5.9.27}$$

Therefore,

$$\langle s_{i+r}s_i \rangle \sim (\lambda_-/\lambda_+)^r. \tag{5.9.28}$$

The above consideration clearly indicates that as long as the transfer matrix is a finite matrix no phase transition exists. If the interaction range is not finite, then the transfer matrix is no more finite-dimensional and there is a possibility that even a 1D system can have phase transitions (as we know for the Kac model in 1D).

Exact solutions are very useful of course, but the reasons for successfully solving a problem exactly may be a rather unimportant peculiarity from the physics point of view. In 2-space the richness of the conformal symmetry group is often a major reason for exact solutions and the reason is not at all trivial or mere peculiarity, but still two dimensional world is an extremely special world. Was the Onsager exact solution for the 2-Ising model decisive for the understanding of phase transition?[76] There are people who say that Onsager's result for the first time demonstrated that the equilibrium statistical mechanics framework could capture phase transition, but Peierls' work was far before the exact solution.

---

[76] Onsager's much greater contribution to statistical physics is his contribution to nonequilibrium theory. This point seems often ignored as can be seen explicitly in S. G. Brush, *Statistical Physics and the Atomic Theory of Matter, from Boyle and Newton to Landau and Onsager* (Princeton UP, 1983).

## 5.10 Real space renormalization

Kadanoff's idea may be summarized as follows:
We introduce some method $B$ to coarse-grain the system. This method also dictates the extent of the spatial scale reduction rate $\ell$. The coarse-graining method $B$ may be understood as a map from a configuration $S$ (this may be a field or spin configuration $\{s_i\}$) of the original system to a configuration of the reduced system. Fig. 5.10.1 illustrates two examples. The important point of $\mathcal{K}$ is that it is a map: given a configuration $S$, $\mathcal{K}(S)$ is unique. However, it is not an injection (one-to-one map), since it is a kind of coarse-graining.

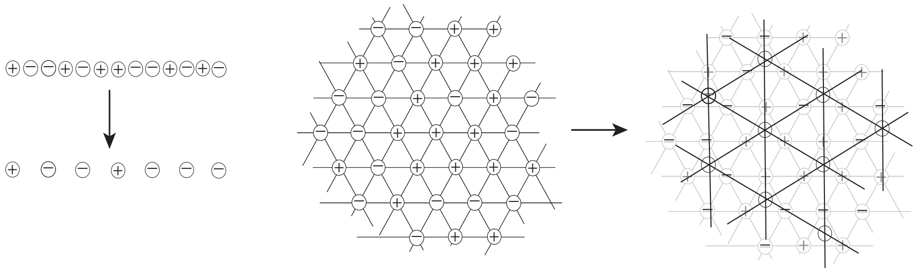


Fig. 5.10.q 1 Left: Decimation of 1-Ising model, $\ell = 2$; Right: Blocking of 3 spins of the triangular lattice 2-Ising model. $\ell = \sqrt{3}$. The value of the block spin is determined by the majority rule: the block spin is up (down) if two or more spins being blocked are up (down).

We wish to compute the partition function

$$Z = \sum_S e^{-H(S)}, \tag{5.10.1}$$

where the microstate (or the microscopic configuration) is denoted as $S$. Here, $\beta = 1/k_BT$ is not explicitly written; as we will see soon, there is no point to separate temperature out, because the coefficients in the coarse-grained (or renormalized) Hamiltonian depend in a complicated way on temperature. Let $\mathcal{K}(S) = S'$, where $S'$ is the coarse-grained configuration, which is a microstate for the coarse-grained system, as illustrated in Fig. 5.10.1. We introduce a generalized Kronecker $\delta$:

$$\Delta(C, C') = \begin{cases} 1, & \text{if } C = C', \\ 0, & \text{otherwise,} \end{cases} \tag{5.10.2}$$

where $C$ and $C'$ are configurations of a given system. Since $\mathcal{K}$ is a map, for any fixed $S$

$$\sum_{S'} \Delta(\mathcal{K}(S), S') = 1. \tag{5.10.3}$$

Therefore, obviously we have

$$Z = \sum_{S} \left[ \sum_{S'} \Delta(\mathcal{K}(S), S') \right] e^{-H(S)}. \tag{5.10.4}$$

Now, we change the order of summations:

$$Z = \sum_{S'} \left[ \sum_{S} \Delta(\mathcal{K}(S), S') e^{-H(S)} \right]. \tag{5.10.5}$$

We may define the Hamiltonian $H(S')$ for the coarse-grained system as

$$H'(S') = -\log \left[ \sum_{S} \Delta(\mathcal{K}(S), S') e^{-H(S)} \right], \tag{5.10.6}$$

which defines the coarse-grained system on the coarse-grained lattice. We may rewrite the partition function as[77]

$$Z = \sum_{S'} e^{-H'(S')}. \tag{5.10.7}$$

If $\mathcal{K}$ blocks adjacent $k$ spins into a single 'block spin,' $H'$ is a spin Hamiltonian for block spins.

In this way $H'$ is constructed from $H$. As the examples in Fig. 5.10.1 illustrate, $H'$ can live on a lattice that is similar to the original lattice. Actually what we have accomplished by $H \to H'$ is a construction of the renormalization group transformation $\mathcal{R}$ with $\ell$ being geometrically easily read off from the figures ($\ell = 2$ for the 1D Ising decimation and $\ell = \sqrt{3}$ for the triangle lattice example).

As a simple example, let us study 1-Ising model with the aid of *decimation*. This procedure thins the spins through summing over a subset of spins, keeping the rest

---

[77] As the reader may have noticed, more generally, $\Delta$ may be replaced with a nonnegative function $P(S, S')$ such that $\sum_{S'} P(S, S') = 1$.

fixed. Look at Fig. 5.10.1(left). The partition function reads (here, $K = \beta J$; we absorb $\beta$ into the original Hamiltonian)

$$Z = \sum_{s,\sigma} \cdots e^{K(s_{-1}\sigma_0+\sigma_0 s_1)} \cdots, \tag{5.10.8}$$

where spins at the even lattice positions are written as $\sigma$. We sum over all $\sigma$ states. The result is a product of the terms of the following form:

$$\sum_{\sigma=\pm 1} e^{K(s_{-1}\sigma_0+\sigma_0 s_1)} = 2\cosh K(s_{-1}+s_1). \tag{5.10.9}$$

Equating this with the form $\propto e^{K' s_{-1}s_1}$, we can fix $K'$ with the aid of the fact that $s^2 = 1$:[78]

$$K' = \frac{1}{2}\log\cosh 2K. \tag{5.10.10}$$

Thus, we have constructed a map from the original Hamiltonian to the coarse-grained Hamiltonian with $\ell = 2$:

$$H = \sum_{i\in \boldsymbol{Z}} K s_i s_{i+1} \to H' = \sum_{i/2\in \boldsymbol{Z}} K' s_i s_{i+2}. \tag{5.10.11}$$

Starting from some positive $K$ and iterating (5.10.10), we see (e.g., graphically) clearly that $K \to K' \to \cdots \to 0$ quickly, consistent to the fact that there is no phase transition for $T > 0$.

Let us study a more interesting example: the triangular lattice Ising model.**q** In contrast to the 1D case, in higher dimensions it is generally the case that coarse-graining produces multi-spin interactions, even if the original model contains only binary spin interactions as in the present example.[79]. However, we wish to be as simple as possible, so we use a (crude but still interesting) approximation that under $\mathcal{K}$ illustrated in Fig. 5.10.1(right), the Hamiltonian preserves its shape (that

---

[78] If one wishes to be more precise, assume $e^{A+K's_{-1}s_1}$. If the spin sum is zero, $2 = e^{A-K'}$; otherwise, $2\cosh(2K) = e^{A+K'}$, so we obtain $e^{2K'} = \cosh(2K)$. The free energy of the original system cannot be determined by $H'$ only, but depends on $A$ as well. In reality, the contribution of $A$ is larger, but it does not have the singularity that is responsible for the critical singularity. The free energy $f_s$ in Kadanoff's theory was just the singular part removing this contribution of $A$.

[79] This is the explanation of why as soon as renormalization transformation is applied the system leaves the natural trajectory in Fig. 5.7.3. Here, we assume that the form of the Hamiltonian does not change. This is equivalent to assuming that the RG flow and the 'natural trajectory of the system' agree. We can see from Fig. 5.7.3 that we study the trajectory going through $H^*$.

is, we assume that the RG flow does not leave the phase diagram. Recall Section 5.7.):[80]

$$H = \sum K s_i s_j + h s_i \to H' = \sum K' s'_\alpha s'_\beta + h' s'_\alpha, \tag{5.10.12}$$

where $s'_\alpha$, etc. denote the block spins defined by the majority rule: if two or more spins are up (down) in the block, the block spin is up (down). Fig. 5.10.2 explains the block spins more explicitly.

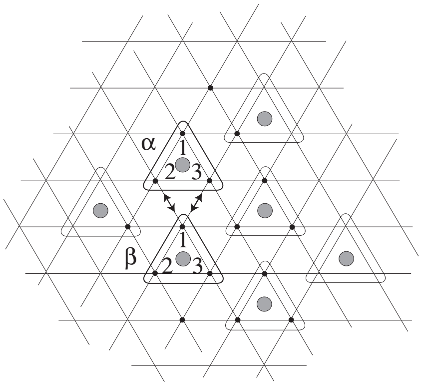


Fig. 5.10.2**q** Triangular lattice and the block spins $\alpha$ and $\beta$. 1, 2, 3 denote the original spins (small black dots). The rounded triangles denote block spins, and gray disks indicate the positions of the block spins.

For simplicity, let us study the small $h$ case; we ignore its effect on the coarse-grained coupling constant. Since we are interested in the macroscopic global behavior of the mode, we need not worry about the intrablock spin interactions.[81] Therefore, the 'block spin $\alpha$'-'block spin $\beta$' interaction energy must be equal to the sum of interaction energies among the original spins belonging to different blocks. Thus, as can be seen from Fig. 5.10.2, we may demand

$$K' s'_\alpha s'_\beta = K(s_{\alpha 2} s_{\beta 1} + s_{\alpha 3} s_{\beta 1}) \tag{5.10.13}$$

on the average (we cannot demand this exactly). That is, the block spin $\alpha$-$\beta$ interaction is supported by two 'actual' interactions: interactions between $\beta 1$ spin and $\alpha 2$ and $\alpha 3$ spins.

If we wish to relate $K$ and $K'$, we must relate $s$ and $s'$. The basic idea is that near the critical point the correlation length $\xi$ is large, so

$$K' s'_\alpha s'_\beta = K(\langle s_{a2}\rangle_{s'_\alpha} \langle s_{\beta 1}\rangle_{s'_\beta} + \langle s_{a3}\rangle_{s'_\alpha} \langle s_{\beta 1}\rangle_{s'_\beta}), \tag{5.10.14}$$

---

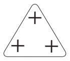
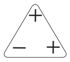
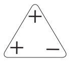
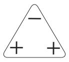
[80] More accurate handling of this problem can be seen in Th. Niemeijer and J. M. J. van Leeuwen, "Wilson theory for spin systems on a triangular lattice," Phys. Rev. Lett. **31**, 1411 (1973).

[81] They shift the origin of the free energy, but it has nothing to do with the correlation length, so they correspond to the non-singular part of the free energy.**r** Recall that we discussed the singular part of the free energy; we are picking up the singular part only.

where $\langle s\rangle_{s'}$ is the average of the original spin $s$ in the block spin whose value is $s'$ (a conditional average), and

$$s'_\alpha = \mathrm{sgn}(\langle s_{\alpha 1}\rangle_{s'_\alpha}). \tag{5.10.15}$$

The following table tells us the original spin configuration compatible with $s'_\alpha = +1$ (i.e., the majority up). The last line in the table is the internal energy of the block spin that determines how a particular internal configuration is likely.

| |  |  |  |  |
|---|---|---|---|---|
| $s_{\alpha 1}$ | +1 | +1 | +1 | −1 |
| energy | $-3K$ | $+K$ | $+K$ | $+K$ |

Therefore, we obtain

$$\langle s_{\alpha 1}\rangle_+ = \phi(K) \equiv \frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}. \tag{5.10.16}$$

By symmetry $\langle s_{\alpha 1}\rangle_- = -\langle s_{\alpha 1}\rangle_+$, so we can write

$$\langle s_{\alpha 1}\rangle_{s'_\alpha} = \phi(K)s'_\alpha. \tag{5.10.17}$$

We have obtained all the ingredients to relate $K$ and $K'$. (5.10.13) now reads

$$K's'_\alpha s'_\beta = 2K\phi(K)^2 s'_a s'_\beta, \tag{5.10.18}$$

or

$$K' = 2K\phi(K)^2. \tag{5.10.19}$$

It is clear that there is no point to separate $k_BT = 1/\beta$ out in the Boltzmann factor; even if $K$ may be written as $\beta J$ for some $T$-independent $J$, the $T$ dependence of $K'$ is not this simple.

Since we have assumed that $h$ is small, we may simply ignore its effect on $K'$, and we require

$$h's'_\alpha = h(s_{\alpha 1}+s_{\alpha 2}+s_{\alpha 3}), \tag{5.10.20}$$

so we immediately obtain

$$h' = 3h\phi(K). \tag{5.10.21}$$

This completes our construction of $\mathcal{R}:\ (K,h)\to(K',h')$. $\ell=\sqrt{3}$.

We must not forget what we are really doing. Look at Fig. 5.7.3 again. We are looking at the flow due to $\mathcal{R}$ in the Hamiltonian space. In the present case we have parameterized the Hamiltonians in terms of $K$ and $h$ (or we have introduced the coordinate system $(K, h)$ to the unstable manifold of the critical point in the Hamiltonian space), so we can describe the Hamiltonian flow in this two dimensional space.[82] For small $h$ we may rely on (5.10.19) and (5.10.21) to study the flow in this space, the $(K, h)$-plane. First, let us look for fixed points $K_F, h_F$ determined by

$$K_F = 2K_F\phi(K_F)^2,\ h_F = 3h_F\phi(K_F). \tag{5.10.22}$$

$K_F = 0$ is certainly a solution, but $\phi = 1/\sqrt{2}$ gives $K^* = (1/4)\log(1 + 2\sqrt{2}) \simeq 0.3356\cdots$. For all $K_F$ $h_F = 0$ is a solution. There is no other finite solution.[83] That is, $(K, h) = (0, 0)$ or $(K^*, 0)$ are the fixed points.[84]

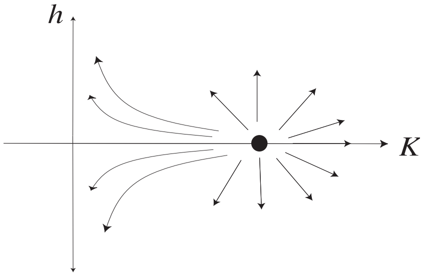


Fig. 5.10.4 The RG flow for the triangular lattice Ising model in the $(K, h)$-plane. The black dot denotes the location of the nontrivial fixed point $(K^*, 0)$. The origin is also a fixed point. This figure corresponds to Fig. 5.7.2; larger $K$ corresponds to lower temperature.

We have already seen that the critical surface is a codimension 2 manifold,[85] and the stable fixed point in it corresponds to the critical point. The critical point has a 2-dimensional unstable manifold. The $(K, h)$-plane we have considered above corresponds to this unstable manifold. Thus, the unstable fixed point $(K^*, 0)$ is the critical point.

In the Kadanoff construction the coordinate system $(\tau, h)$ spans this manifold. As can be seen from $\mathcal{R}\tau = \tau\ell^{y_1}$, we must study the local behavior of $\mathcal{R}$ near the critical fixed point (see Fig. 5.7.3 again).[r] That is, we have only to study the linearization of $\mathcal{R}$ around the critical point Hamiltonian $H^*$. Let the deviation of Hamiltonian from $H^*$ be $\delta H$.[q] Then, we study

$$\mathcal{R}(H^* + \delta H) = \mathcal{R}H^* + D\mathcal{R}\delta H = H^* + D\mathcal{R}\delta H, \tag{5.10.23}$$

where $D\mathcal{R}$ denotes the derivative (or the Jacobi matrix) of $\mathcal{R}$ at the fixed point $H^*$.

---

[82] Actually, this 2-space is an approximation of the unstable manifold of $H^*$ (the orthogonal complement of the critical surface) as will be discussed soon.

[83] Don't divide the equation with zero.

[84] $K_F = \infty$ is a fixed point corresponding to the ordered phases or $T = 0$.

[85] That is, there are two directions that are not contained in this surface, or there are a plane (= a two-dimensional object) perpendicular to the tangent to this surface.

That is, we study the linearized flow near the fixed point:

$$\delta H' = D\mathcal{R}(H^*)\delta H. \tag{5.10.24}$$

**The eigenvalues of $D\mathcal{R}$ is the power of $\ell$**
Let us explicitly write the scaling factor $\ell$ for the renormalization transformation as $\mathcal{R}_\ell$. A very important property of $\mathcal{R}_\ell$ is

$$\mathcal{R}_\ell \mathcal{R}_{\ell'} = \mathcal{R}_{\ell\ell'}. \tag{5.10.25}$$

This implies (the chain rule; cf. $d(f(g(x))/dx = f'(g(x))g'(x))$

$$D\mathcal{R}_{\ell\ell'}(H) = D\mathcal{R}_\ell(\mathcal{R}_{\ell'}(H))D\mathcal{R}_{\ell'}(H), \tag{5.10.26}$$

so at the fixed point

$$D\mathcal{R}_{\ell\ell'}(H^*) = D\mathcal{R}_\ell(H^*)D\mathcal{R}_{\ell'}(H^*). \tag{5.10.27}$$

Let $\lambda(\ell)$ be an eigenvalue of $D\mathcal{R}_\ell(H^*)$. Then, (5.10.27) implies

$$\lambda(\ell\ell') = \lambda(\ell)\lambda(\ell'). \tag{5.10.28}$$

Since we may assume that $\lambda(\ell)$ is a continuous function of $\ell$, (5.10.28) implies that

$$\lambda(\ell) = \ell^\alpha. \tag{5.10.29}$$

In this way, RG completely justifies the Kadanoff construction near the critical fixed point.

Now, let us complete our study of the triangular lattice Ising model, computing the critical exponent $\nu$. Our map $\mathcal{R}_\ell : (K, h) \to (K', h')$ with $\ell = \sqrt{3}$ given by (5.10.19) and (5.10.21) can be linearized around the nontrivial fixed point $(K^*, 0)$ (the parametric representation of $H^*$) as (derivatives are evaluated at the nontrivial fixed point)

$$D\mathcal{R}_\ell(K^*, 0) = \begin{pmatrix} \frac{\partial K'}{\partial K} & \frac{\partial h'}{\partial K} \\ \frac{\partial K'}{\partial h} & \frac{\partial h'}{\partial h} \end{pmatrix}. \tag{5.10.30}$$

Since in our approximation $K'$ is not affected by $h$, the eigenvalues are simply read off as

$$\lambda_1 = \frac{dK'}{dK} = 1.634\cdots, \ \lambda_2 = \frac{\partial h'}{\partial h} = 3\phi(K^*) = \frac{3}{\sqrt{2}}. \tag{5.10.31}$$

The eigenvalue related to the temperature is $\lambda_1$. Our general discussion tells us that $y_1$ is related to this as $\lambda_1 = \ell^{y_1}$, where $\ell = \sqrt{3}$ (the increase of lattice spacing by blocking spins) in our example. Therefore,

$$y_1 = \log 1.634/ \log \sqrt{3} \simeq 0.8939\cdots, \tag{5.10.32}$$

which is $1/\nu$, so $\nu = 1.1187\cdots$. The exact value is 1. The reader may think the result is not impressive, but the mean field theory gives 1/2. Incidentally, $y_2 = \log(3/\sqrt{2})/\log\sqrt{3} = 1.368\cdots$. The exact value is $y_2 = (\gamma+\beta)/\nu = 15/8 = 1.875$, and the mean field theory gives 3.[86]

## 5.11 Field theoretical model and renormalization

Real space renormalization group procedures coarse-grain the spatial configurations, so we can expect that continuous field theory should be enough to describe the second order phase transition. There are several approaches to construct field theoretical models from the usual microscopic models, but none of them is well controlled.[87] Therefore, for practical physicists the best way is to write it down using her intuition and some general constraints such as symmetry.[88]

We write the Hamiltonian in the following form:

$$H = \int d^d\boldsymbol{r}\mathcal{L}, \tag{5.11.1}$$

where $\mathcal{L}$ is the Hamiltonian density[89] that is a local function of the '*order parameter field*' $\phi$ and its spatial derivatives. As an order parameter here we choose naively something that is zero in disordered phases and nonzero otherwise. The magnetization density is a respectable order parameter for magnets.[90]

First of all, we are interested in the systems for which spatially uniform phases away from the phase transition temperature are stable. For such systems nonzero

---

[86] $\beta = 0.704$ (the exact value is 1/8= 0.125; the mean field value is 1/2), $\gamma = 0.827$ (the exact value is 1.75; the mean field value is 1). $\alpha = -0.236$ (the exact behavior is logarithmic singularity; the mean field singularity is discontinuity).

[87] It is generally believed that making the renormalization group method rigorous is the only way to control this.

[88] That is, use of *abduction* in the sense of Peirce is the most appropriate. See Y. Oono, *The nonlinear world* (Springer, 2012) Chapter 4.

[89] The symbol $\mathcal{L}$ is used because this is the (Euclidean version of ) the Lagrangian density in the ordinary quantized field theory.

[90] The order parameter field may be (appropriately normalized) number density for fluids, or (appropriately normalized) concentration difference for binary mixtures.

gradient of the field should be energetically penalized. Thus, $\mathcal{L}$ should at least contain the term like

$$\mathcal{L} = a(\nabla\phi)^2 + \cdots, \tag{5.11.2}$$

where $a$ is a positive constant.[91] Here, we have assumed that our system is macroscopically isotropic and uniform = rotationally and translationally symmetric. We must not forget at what length scale we wish to use the continuum field model. We wish to perform statistical mechanics (statistical field theory) starting with the Hamiltonian we are constructing. Therefore, we are seeking for a description at the 'mesoscopic scale.' That is, the minimum discernible scale is much larger than the lattice spacing, but is much shorter than the correlation length near the critical point. The nonlinear behaviors we observe macroscopically are due to large scale fluctuations, so they should not exist at our mesoscopic Hamiltonian level. Therefore, we expect that the restoring dynamics is linear. Now, as the meaning of the Laplacian discussed before tells us that the leading order contribution must be a quadratic form of the gradient.

If the temperature is high ($T > T_c$), the local average of $\phi$ should be zero. That is, $\mathcal{L}$ must favor $\phi = 0$. The easiest way to realize this is to add $\mathcal{L}$ a term proportional to $\phi^2$:

$$\mathcal{L} = a(\nabla\phi)^2 + b\phi^2 + \cdots, \tag{5.11.3}$$

where $b > 0$.[92] If $T < T_c$, $\phi \neq 0$ must be encouraged. That is, $\phi = 0$ must be energetically penalized. The easiest way is to flip the sign of $b$ in (5.11.3). However, then indefinitely large $|\phi|$ is encouraged, and the system becomes unstable. To stabilize the system or to confine $\phi$ in a finite range, we must penalize too large $\phi$. The easiest way is to add a term proportional to $\phi^4$:

$$\mathcal{L} = a(\nabla\phi)^2 + b\phi^2 + c\phi^4, \tag{5.11.4}$$

where $c > 0$. We can change the space length unit, and also scale the field appropriately,[93] so we may set $a = 1/2$ without any loss of generality. Thus, we have arrived at a model defined by the Hamiltonian:

$$H = \int d^d\boldsymbol{r} \left[\frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}\tau_0\phi^2 + \frac{g_0}{4}\phi^4\right], \tag{5.11.5}$$

---

[91] $(\nabla\phi)^2$ implies the scalar product $= (\partial_1\phi)^2 + (\partial_2\phi)^2 + \cdots$; this is a standard abuse of symbols in physics.

[92] Again, we wish to have a linear equation for small fluctuations, so the lowest order term is quadratic in $\phi$. Although this is a plausible motivation, the final judge of any model is Nature. Therefore, a sensible strategy of model building is to make the simplest and mathematically best behaved model, and check its outcome. If it does not work, let us make a more sophisticated model.

[93] $\phi$ is the variable over which we take sum when we compute the partition function, so it is a dummy variable.

where $\tau_0$ and $g_0(> 0)$ are constants. This is the celebrated $\phi^4_d$ *model.* Such continuous Hamiltonians defined by simple polynomials are generically called the *Ginzburg-Landau Hamiltonian.*

To understand the importance of the parameters,[94] let us perform dimensional analysis of the model. The key observation is that $H$ is dimensionless:[95]

$$[H] = 1. \tag{5.11.6}$$

Since integration is dimensionally equivalent to multiplying length, and since differentiation is equivalent to dividing with length, we have

$$\left[\int d^d\boldsymbol{r}(\nabla\phi)^2\right] = L^d([\phi]/L)^2 = 1. \tag{5.11.7}$$

Here, $L$ is the dimension of length. From this, we obtain

$$[\phi] = L^{1-d/2}. \tag{5.11.8}$$

Using this and $[H] = 1$, we obtain

$$\left[\int d^d\boldsymbol{r}\tau_0\phi^2\right] = [\tau_0]L^d(L^{-1d/2})^2 = 1 \Rightarrow [\tau_0] = L^{-2}. \tag{5.11.9}$$

The result may be read off immediately, if $[\tau_0] = [\nabla^2]$ is recognized by comparing the two terms in the Hamiltonian. Using this approach, we immediately obtain $[g_0\phi^2] = [\tau_0]$, or

$$[g_0] = L^{d-4} = L^{-\varepsilon}, \tag{5.11.10}$$

where $\varepsilon = 4 - d$, famous for the $\varepsilon$-expansion.

Suppose we are interested in phenomena observable at the scale $L$. Then, whether a parameter is important or not may be discussed in terms of dimensionless parameters containing the parameters and $L$. $\tau_0L^2$ is such a dimensionless parameter. This is large, if $L$ is large. That is, if we are interested in macroscopic observables, $\tau_0$ is important. $g_0L^\varepsilon$ is also dimensionless, so if the spatial dimensionality is less than 4, we must respect $g_0$.[96] These parameters are called *relevant parameters*. By the way,

---

[94] Do not forget the field is a dummy variable that are integrated out, so the parameters are the only physically meaningful quantities.

[95] because the probability should not change even if we change units.

[96] Of course, we must if $\tau_0 < 0$, but what is meant here is that even if $\tau_0 > 0$.

if a parameter is relevant, its effect is decisive, but as we have just seen, if $\varepsilon$ is very small, then it is almost *marginal* (that is, dimensionless or in the present context, $L$-independent) and the nonlinear effect may be treated perturbatively. This is the idea of the *$\varepsilon$-expansion method*.[97] Here, 'perturbatively' does not imply that the perturbation series converge. They should not, if the effect of $g_0$ is *interesting*, that is, causes qualitative differences. Thus, the perturbation series are divergent, and are at best asymptotic.[98]

We can also make a dimensionless parameter $g_0\tau_0^{-\varepsilon/2}$. If we are close to the critical point this quantity becomes large in 3-space. That is, we cannot ignore the nonlinear term $\propto \phi^4$. In other words, we must respect large fluctuations as we know well already. Thus, for $g_0\tau_0^{-\varepsilon/2} > 1$, the mean field theory becomes unreliable. This criterion is called the *Ginzburg criterion*.

Our 'derivation' of the field theory above looks quite arbitrary. Therefore, it is natural to consider a more general Hamiltonian than the $\phi^4_d$-model (the suffix $d$ means the spatial dimensionality) called $P(\phi)_d$-model ($P$ is for 'polynomial') with the Hamiltonian density $\mathcal{L}$ given by

$$\mathcal{L} = \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}\tau_0\phi^2 + \frac{1}{4}g_0\phi^4 + \cdots + \frac{1}{2n}u_{2n}\phi^{2n}. \tag{5.11.11}$$

For example, let us look at $u_6$. We have

$$1 = [u_6][\phi]^6 L^d = [u_6]L^{6-3d} \;\Rightarrow\; [u_6] = L^{2d-6}. \tag{5.11.12}$$

For $d = 3$ this is marginal and is irrelevant for $d > 3$. We cannot always ignore it in 3-space but the correction is not large.[99]

**Building up of 'low power terms' from higher power terms**

The reader might have thought, if $g_0 = u_6 = 0$ and $\tau_0 > 0$, then we need not worry about 'nonlinear terms.' However, do not forget that, for example, the $\phi^8$ term looked 'from distance' (i.e., after renormalization) looks like the $\phi^4$ term. If the Feynman diagram for the $\phi^{2n}$ term is drawn, it is easy to see that it expresses a 'collision' (nonlinear interaction) effect among $n$ harmonic components (modes) of fluctuations. Look at the 'collision' among two high (spatial) frequency modes and two low frequency modes from distance. It may look like a collision of two low

---

[97] K. G. Wilson and M. E. Fisher, "Critical exponents in 3.99 dimensions," Phys. Rev. Lett. **38**, 240 (1972).

[98] **Warning** Divergence and asymptoticity do not have any logical relation; an asymptotic series need not be divergent nor a divergent series asymptotic.

[99] This is effective only when the $\phi^4$ term cannot stabilize the system.

frequency modes locally modulated by high-frequency modes, so it is effectively a two body collisions. Thus, the model with $\mathcal{L} = (1/2)(\nabla\phi)^2 + \tau_0\phi^2 + u_{100}\phi^{100}$ describes phenomenologically the same phenomenon as the $\phi^4$ model; they are in the same 'universality class.'

It should also be clear to the reader that the $\phi^4$ term modifies the $\phi^2$ term. This implies that to describe the critical phenomenon, $\tau_0 < 0$ must be carefully chosen.

How about higher order spatial derivatives of $\phi$? For example, we could imagine a term $a(\nabla\phi)^4$. Its dimension is $[a]L^{-2d}$, so we may ignore it when we are interested in large scale phenomena.[100] Such parameters are called *irrelevant* parameters. We can easily see the higher order derivative terms such as $(\Delta\phi)^2$ can be ignored.

We may conclude that the $\phi^4_d$-model is enough to describe the critical phenomenon.

If we wish to make a continuum field theory without cutoff $\ell$ (or more commonly, without the cutoff momentum $\Lambda$), the dimensionless parameter of interest is $\Lambda^{-\{2n-(n-1)d\}}u_{2n}$. If this quantity diverges in the $\Lambda \to \infty$ limit, the operator $\phi^{2n}$ is said to be *non-renormalizable* in field theory (or in high energy physics). The theory with such a term is not well defined as a continuum field theory, because we cannot take the continuum limit $\Lambda \to \infty$. If $\Lambda^{-\{2n-(n-1)d\}}u_{2n}$ vanishes or finite in the $\Lambda \to \infty$ limit, we say the operator is *renormalizable*.

Notice that nonrenormalizable quantities correspond to irrelevant quantities (see Fig 5.10.1).

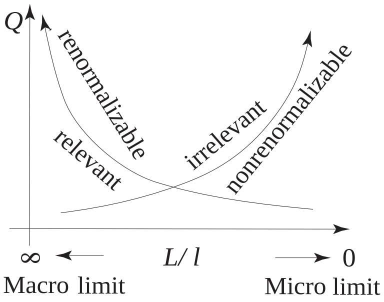


Fig. 5.11.1 Irrelevant vs. relevant or nonrenormalizable vs. renormalizable effects. Here $L$ is the length scale and $\ell$ is the cutoff length scale ($\ell \sim 1/\Lambda$).

If the effect of a certain microscopic term converges in the macroscopic limit to the same limit, then we may ignore it from the theory, because it does not give different results for macro-observables of different systems. If the macroscopic limit does not converge, it means that a small change in the term must have serious effects on the macroscopic observables. Therefore, its effect must be subsumed into a macroscopic parameter (e.g., materials constant) that alters macroscopic observables.

---

[100] Of course, we ignore the 'building up' of lower order powers as explained above.

We have come to a conclusion: if we wish to study macroscopic universal behaviors of a system, we have only to take into account the renormalizable terms when we make a model. Nonrenormalizable terms can simply be ignored. Therefore, a field theory that explains our low energy world can be assumed to be renormalizable. That is, renormalizability may be used as a criterion for a respectable theory. Emerging from this conclusion is a view that all the field theories are phenomenological theories.[101]

What actually does renormalization group (RG) do? Or, what can we obtain by RG? Let us look at a typical result. For $T > T_c$, the correlation length behaves as

$$\xi = A(T - T_c)^{-\nu}. \tag{5.11.13}$$

Here, $T$ and $\xi$ are observables; $A$ and $T_c$ are materials constants; $\nu$ or the functional structure $x^{-\nu}$ is the universal structure. This universal structure is for a class of phenomena or materials (e.g., sharing the same spatial and spin dimensionality). Another example is the size of a polymer chain in a solvent: its mean square end-to-end distance reads

$$\langle R^2 \rangle = AM^{2\nu}, \tag{5.11.14}$$

where the LHS and the molecular weight $M$ are observables, $A$ is a materials constant (depending on the polymer-solvent pair), and $x^{2\nu}$ is the universal functional form. This holds for any pair of a sufficiently long polymer chain and a solvent that dissolves the polymer well. Thus, we see the following general structure in the RG results:

$$\text{Phenomenology} = \text{Universal Structure} + \text{Materials Constants}. \tag{5.11.15}$$

Here, "*Phenomenology*" means a general description of a class of phenomena, and "Materials Constants" means something dependent on the individual idiosyncrasy of each system; the fetish part in Fisher's terminology. We call a collection of phenomena or systems sharing the same universal structure a *universality class*. The concept of universality was consciously introduced by Kadanoff.[102] From this point of view, the principal aim of statistical physics is to identify universality classes and

---

[101] The following review article is strongly recommend: K.-I. Aoki, "Introduction to the non-perturbative renormalization group and its recent applications," Int. J. Mod. Phys. B **14**, 1249-1326 (2000).

[102] He once told the author that when he started to study critical phenomena, he assumed everything was universal; then by and by it turned out to be known that there are different universality classes.

study their universal features.

The aim of the previous section was to explain a method to extract universal features of a system. The guiding idea is to observe a system from distance with fixed eyesight; what continues to be observed should be universal. The strategy combining coarse-graining and shrinking is called the *Wilson-Kadanoff renormalization group method*. On the other hand, there is an approached based on our observation (5.11.15); if we alter microscopic details (this could be changing materials) the second term of (5.11.15) changes. This change is often very sensitive to the alteration (see Fig. 5.11.1), if we could isolate microscopic-detail sensitive parts, the rest should be selected as the universal features. The renormalization group method based on this idea is called the *Stückelberg-Petermann renormalization group method*.[103]

## 5.12 Spontaneous symmetry breaking

**r**If an equilibrium state (precisely, a pure Gibbs state) has a symmetry group which is a genuine subgroup of the symmetry group of the system Hamiltonian, we say the symmetry is *spontaneously broken*. Certainly, the symmetry is spontaneously broken below $T_c$ for 2-Ising model. In this case the symmetry that is broken is described by a discrete group (up-down symmetry).[104] Really interesting cases of spontaneous symmetry breaking occur if the broken symmetry is continuous such as rotation.[105]

Consider a Heisenberg magnet as an example (cf. Fig. 5.12.1).

---

[103] This is the so-called field-theoretical renormalization group theory, and was invented in the 1940s. However, it is fair to say that its true meaning was not understood until the advent of the Wilson-Kadanoff renormalization group method. As an introductory textbook, M. Le Bellac, *Quantum and Statistical Field Theory* (Oxford 1991) may be recommended. Field theoretical renormalization is technically rather complicated, but the basic idea can be understood without technical difficulty as is explained with the aid of the von Koch curve. This and a sample calculation of field theoretical renormalization group calculation ($\varepsilon$-expansion) can be found in Supplementary Pages. See also Chapter 3 of Y. Oono, *The Nonlinear World* (Springer, 2012).

[104] $Z_2$ group.

[105] $O_3$.

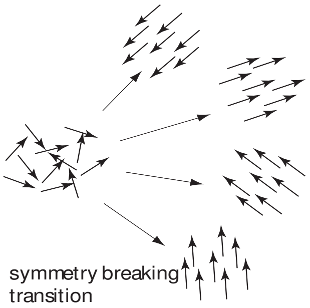


Fig. 5.12.1 Symmetry breaking results in an ensemble of symmetry broken phases collectively representing the whole symmetry of the system. (This illustration corresponds to a transition from a paramagnetic phase to a ferromagnetic phase.)

When a spontaneous breaking of a continuous symmetry occurs, there are two important consequences, development of collective excitations called the *Nambu-Goldstone bosons* (NG bosons) and the (generalized) *rigidity* in the ordered phase. The NG bosons refer to long wave length collective excitations in the ordered phase (like acoustic phonons = sound waves in solids) whose excitation energy tends to zero in the long-wavelength limit. Rigidity implies that the change of the order parameter in a small part of the system propagates throughout the system (like the rigidity of a solid body: if one end is twisted, the other end follows).[106]

The reason for these consequences is easy to understand, if we pay due attention to the nature of breaking of the continuous symmetry. All possible symmetry broken phases (see Fig. 5.14.1) have the same energy, because they can be transformed into each other with an element of the symmetry group of the system Hamiltonian. A local deformation is energetically unfavorable, so the system changes to a new phase compatible with the local deformation to remove the excess deformation energy. This change is observed as the manifestation of rigidity. The energy cost per unit cross-sectional area of deformation of the order parameter field along an axis (say, $x$-axis) may be estimated as

$$\int_0^L \left(\frac{\partial \phi}{\partial x}\right)^2 dx \simeq \left(\frac{\delta\phi}{L}\right)^2 L = \frac{(\delta\phi)^2}{L}, \tag{5.12.1}$$

where $\delta\phi$ is the overall change of the order parameter. This implies that if $L$ (the total length) is large enough, the total energy required to deform by $\delta\phi$ can be made indefinitely small. Thus, excitation of long-wavelength deformations is energetically

---

[106] The change in a 'small part' is in this case kept by an external means. Then, the change eventually propagates to the whole system (i.e., any indefinitely large finite domain follows). Notice that an equilibrium state is stable under any perturbation of any finite domain, if the perturbation is left unconstrained. Do not mix up these different situations.

without any cost, the existence of the NG bosons.

A more explicit illustration with the aid of the spin system (or the $\phi^4$ system) with spin dimension 3 (the Heisenberg model) is as follows. Suppose the spin density $\boldsymbol{\phi}$ (a 3-vector) is governed by the following Hamiltonian density:[107]

$$\mathcal{L} = \frac{1}{2}(\nabla\boldsymbol{\phi})^2 + V(\boldsymbol{\phi}), \tag{5.12.2}$$

and a spontaneously broken symmetry phase is described by $\boldsymbol{\phi} = \boldsymbol{e}_3$ (the magnetization is in the third coordinate direction; we assume that $V(\boldsymbol{e}_3) = 0$ is a minimum; we assume a 'Mexican hat shaped' $V$: Fig. 5.14.2).

Let us expand $\mathcal{L}$ around this, writing $\boldsymbol{\phi} = \boldsymbol{e}_3 + \boldsymbol{\theta}$:

$$\mathcal{L} = \frac{1}{2}(\nabla\boldsymbol{\theta})^2 + \frac{1}{2}V_{33}\theta_3^2 + \cdots. \tag{5.12.3}$$

Due to the symmetry of the potential there is no restoring potential in the directions perpendicular to the magnetization (order). Therefore, say, the first component is governed by

$$\frac{1}{2}(\nabla\theta_1)^2, \tag{5.12.4}$$

so the dispersion relation for, e.g., diffusive modes is $\omega \propto k^2$, that is, long-wave length excitation costs very little. This is an illustration of the NG excitations.

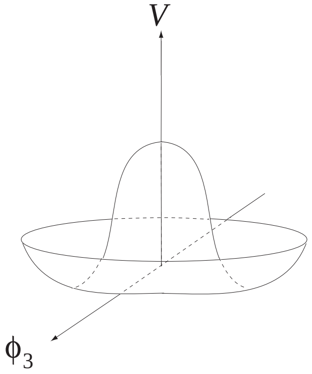


Fig. 5.12.2 The free energy (Landau potential) under symmetry breaking. The potential is rotationally symmetric, but in the ordered phase, the magnetization points in a particular direction. Suppose it is the third coordinate direction. The third component has a restoring force to keep it nonzero, but for the remaining components there is no restoring force along the circular trough.

As can be seen from (5.12.4), the correlation of the displacement in the $x$-direction reads

$$\langle\theta_1(r)\theta_1(0)\rangle \sim \frac{1}{r^{d-2}} \ (r \to \infty). \tag{5.12.5}$$

---

[107] $(\nabla\boldsymbol{\phi})^2$ implies $(\nabla\phi_1)^2 + (\nabla\phi_2)^2 + (\nabla\phi_3)^2$.

(for $d = 2$ this is $\log r$.) This is an alarming result, because it suggests a peculiar behavior in $d = 2$. Indeed, this is the case: there is no symmetry broken phase in $d = 2$ for $n \geq 2$ systems (the *Mermin-Wagner theorem*).[108]

If there is a long range interaction (as in plasmas), however spatially-gently one changes the order parameter, one cannot lower the needed energy cost, because significantly different orientations interact directly through space. Therefore, the NG boson is not expected to exist. Indeed, in plasmas, the so-called plasma oscillation (plasmon) has a lower energy cutoff.

We can summarize representative examples.[109]

| | solid | Heisenberg ferro | superfluid |
|---|---|---|---|
| Broken Symmetry | 3D translational | rotational | phase |
| Order | 3D periodicity | ferromagnetism | superfluidity |
| NG boson | acoustic phonons | spin wave | second sound |
| Rigidity | rigidity | ferromagnetism | superfluidity |

If the system is finite, there is no symmetry breaking.[110] Fig. 5.12.1 implies the following difficulty: if we compute the partition function of a system as usual

$$Z = \sum_{\Gamma} e^{-H}, \tag{5.12.6}$$

because the sum is over all the possible microscopic states, the resultant $Z$ or the free energy of the system is completely symmetric, that is, its symmetry group is identical to that of the microscopic Hamiltonian. This statement is true if the system is finite, because the sum is a finite sum. Thus, taking the thermodynamic limit and introducing Gibbs states are absolutely needed to make a rational and simple framework to understand spontaneous symmetry breaking.

When the intrinsic symmetry is broken, how is a particular phase selected in the

---

[108] It asserts that there is no ordering for 2D Heisenberg model. This is readily extended to include the 2D XY model. The theorem is much harder to prove than the Lee-Yang circle theorem (see IRG Lecture 30 for a proof).

[109] About this section, a strongly recommended reference is: P. W. Anderson, *Basic Notions of Condensed Matter Physics* (Westview Press 1984, 1997), Chapter 2.

[110] To state more practically, the state with a broken symmetry has a life time. For example, for a very small crystal, thermal fluctuation could spontaneously rearrange the crystal axes. Needless to say, if a crystal is not very small such fluctuations occur only very rarely. The agreement of its behavior to the behavior in the thermodynamic limit is practically perfect, because the life-time of a give orientation is very long. However, mathematically, or theoretically, it is still not a true equilibrium state, so thermodynamic limit is taken.

real world? This is selected by extremely small fortuitous external effects or even without such effects by intrinsic thermal fluctuations. If there is a weak external field, the system would react very sensitively to it. Therefore, if one wishes to study a particular phase with the aid of statistical mechanics an appropriate weak field conjugate to the order parameter is introduced to the system Hamiltonian to select the phase. After computing its thermodynamic limit, the field is set to zero. This limit must be take after the thermodynamic limit; if taken zero before the thermodynamic limit, the symmetry breaking field effect disappears. Symmetry breaking means that the thermodynamic limit and the conjugate-field zero limit are not commutative.

## 5.13 First-order phase transition

The most fundamental macroscopic description of a macrosystem in equilibrium is in terms of thermodynamic coordinates. The entropy as a function of the thermodynamic coordinates gives the most complete thermodynamic description of the system. In other words, if we know the internal energy as a function of entropy and work coordinates as $E = E(S, V, \cdots)$, we have a complete thermodynamic description of the system. Generally, it was mentioned that phase transition is characterized by a singularity in a certain thermodynamic potential. In terms of internal energy, a phase transition occurs where the convex function $E = E(S, V, \cdots)$ loses its smoothness. Here 'smoothness' implies the holomorphy as a multivariable function. Since a convex function is continuous, $E$ cannot have any jump. Furthermore, as we see from the Gibbs relation, its continuous differentiability must be satisfied in the region of thermodynamic space meaningful to the system. Thus, internal energy must be a $C^1$ convex function of entropy and work coordinates. Consequently, the worst singularity is the loss of second differentiability. For example, the constant volume specific heat can become not definable. We know at the critical point this indeed happens.

If a $C^1$ convex function loses twice differentiability, what can happen? Let us look at one variable $S$ of $E$. Let us assume that work coordinates (such as the volume) are kept constant. Here we pay attention to the case in which the singularity is isolated. We will not discuss more general cases. Fig. 5.13.1 illustrates $E$ as a function of $S$. The slope of this curve is temperature $T$. Something happening to the second derivative implies that the temperature derivative of $S$ (the constant volume specific heat) has a singularity.

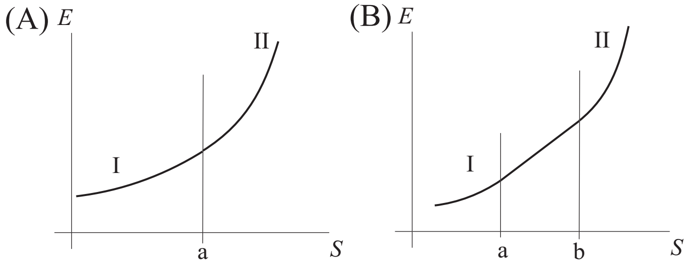


Fig. 5.13.1. When twice differentiability is lost: In (A) it is assume that the second differentiability has a problem at a single point a. In (B) this happens at two points a and b.

In (A) phase I and phase II have the same extensive variable values (the values of thermodynamic densities) at a, so these two phases do not coexist. In this case the order parameter may change continuously. In contrast, in (B) phase I and phase II coexist at a temperature $T$ (= given by the slope of the straight portion between a and b). These two phases are distinct and have different densities. If some density changes discontinuously at the phase transition, it is called a first order phase transition. Otherwise, it is generally called a higher order (usually second order) phase transition; if two phases can coexist, the transition is first order. This happens for (B) (however, even if the transition is first order, phases may not coexist; recall the 2D Ising model). In case (A) a crude sketch of the energy function cannot tell whether the transition is first order or higher.

rTo understand the coexistence of two phases under constant temperature discusssed above, it is convenient to use the thermodynamic potential one of whose independent variables is temperature, that is, the Helmholtz free energy. It is obtained by the Legendre transformation with respect to entropy. We have already seen a general introduction to convex analysis. Here, let us see some detail when there is a phase transition. We know $A = \inf_S[E - ST]$. If this is rewritten in the form standard to convex analysis, it reads $-A = \sup_S[ST - E]$ (i.e., $E^* = -A$). Thus, the free energy is convex upward as a function of temperature (In Fig. 5.13.2 the convex function $-A$ is illustrated).

If a first order phase transition happens and if two phases can coexist, there is a 'linear' portion in the graph of internal energy. This is mapped to a point by the Legendre transformation (Fig. 5.13.2). As can be seen from this, when two phases coexist, thermodynamic states that can be distinguished by thermodynamic coordinates (intuitively, the states distinguishable by different ratios of two phases) are identified and mapped to a single point by the Legendre transformation. In the example illustrated in Fig. 5.13.1 phases with various entropy densities are mapped to a point p designated by a single temperature $T_p$ (the *coexistence temperature*). In

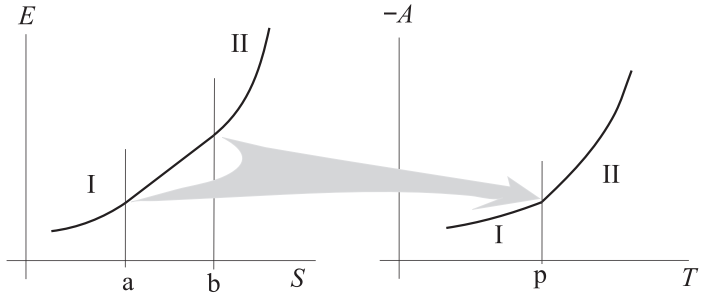


Fig. 5.13.2. **Left** is just the same as Fig. 5.13.1 (B) and depicts $E$as a continuously differentiable function of $S$. $E$ is linear between a and b, and the slopes at a and at b agree with the slope of the linear portion. Phase I occupies left of a, and phase II right of b,and the linear portion describes the coexistence of these phases. The slope of the linear portion is the coexistence temperature $T_p$, corresponding to the break point p of the free energy graph on the right. All the coexisting phases between a and b are mapped to a point p by the Legendre transformation.

this sense, Legendre transformations are not generally one to one transformation. However, it should be noted that from the right graph in Fig. 5.13.2, we can completely reconstruct internal energy as a function of thermodynamic coordinates by the inverse Legendre transformation $E = \sup_T[ST - (-A)]$ (i.e., $E^{**} = E$). This is the implication of the ensemble equivalence.

Up to this point we looked at the phase transition in terms of entropy or temperature, but we already know that phase diagram can depend on other variables as well. For example, we know the phase diagram exhibiting the solid-liquid-gas phase diagram for a system consisting of simple molecules whose coordinates are $T$ and $P$. To glance at a more complicate situation, let us discuss the complete thermodynamic description of this system in terms of internal energy $E = E(S, V)$, and its relation to the description in terms of the Gibbs free energy $G = G(T, P)$, which is obtained by the Legendre transformation of $E$ with respect to $S$ and $V$.

If we draw the graph of internal energy as a function of $S$ and $V$, there are regions surrounded by the 'edges' where second differentiation is not possible. First order phase transitions occur in such regions. Fig. 5.13.3 is the graph of internal energy. The regions designated by various 'coexistence' are ruled surfaces, and the line segments connecting two coexisting phases lie in such regions. The flat triangle marked with t corresponds to the triple point. Generally speaking, flat (hyper) surfaces of the dimension equal to or less than the number of independent variables (the line segments in the ruled surfaces in the above example) correspond to phase coexistence, and at each point in such surfaces the ratio of the amounts of coexisting phases are

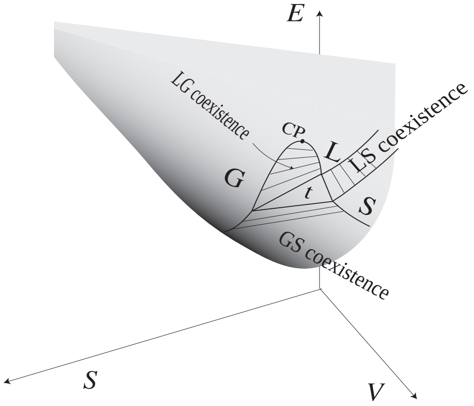


Fig. 5.13.3 The convex function $E = E(S, V)$ looked up from below. Left is its rough sketch of the internal energy 'surface' relief. It is only conceptual and not very accurate. Right is the graph of $E$ looked up from below in the direction of the gray arrow in Left. CP is the gas-liquid critical point. The triangle (which is a flat surface) with t corresponds to the triple point. The surfaces extended from the edges of the triangle are not flat but they are all ruled surfaces; several line segments are drawn in the LS (liquid-solid) coexistence domain to show this. These line segments correspond to the linear portion in Fig. 5.13.2 Left.

uniquely determined.[111] The result of the Legendre transformation that uses $T, P$ as independent variables is illustrated in Fig. 5.13.4.

To study phase transitions, we should know what phases we can observe beforehand. However, this is not an easy problem. In many cases experiments tell us the existence of interesting phases, and then theories are constructed to explain their nature and required microscopic interactions. Then, suppose we know existing phases. How can we know what types of phase transitions we can expect? In many cases mean field approaches are used heuristically, but as has already been warned, mean field theories do not even have a guarantee to predict the existence of the phase transitions. If a model is given in terms of the Ginzburg-Landau Hamiltonian, there is a heuristic way that is often useful, which will be discussed in the next section. The rest of this section looks at some topics relevant to the general features of first order phase transitions.

---

[111] In the above example a triangular flat surface corresponding to the triple point. In this triangle, the ratio of solid:gas:liquid is unique at each point. If there were a flat square, four vertex points are not needed to uniquely specify the position on the square. Thus, the amounts of four phases corresponding to the vertices to a point in the square are not uniquely determined. That is, the phase rule is violated. However, it seems to be a difficult problem to know the microscopic condition for the phase rule to hold.

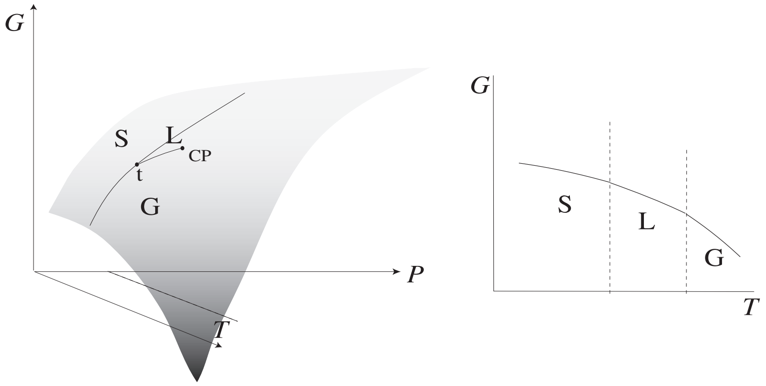


Fig. 5.13.4 Left: A convex upward function $G = G(T, P)$ (only its part is shown). CP is the gas-liquid critical point, and t corresponds to the triple point. The entered curve is the phase coexistence line, which has kinks. The relation between these kinks and the ruled surface in Fig. 5.13.3 is just parallel to that shown in Fig. 5.13.2. Right: the free energy change as a function of temperature under constant pressure.

We have already discussed that a first order phase transition occurs if the cooperation among microscopic elements in the system is not sufficiently strong. Let us see this explicitly within the mean field approximation. For a magnet, suppose that if the (magnitude of the) magnetization per spin $m$ becomes smaller, $J$ in (5.8.8) decreases as illustrated in Fig. 5.13.5.

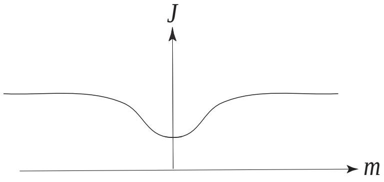


Fig. 5.13.5 Order-dependent coupling constant that induces a first order phase transition. If the order parameter becomes small, the spin-spin interaction becomes weak. In such a model the order would precipitously decreases.

The self-consistent equation for the mean field now reads, instead of (5.8.8),

$$x = 2d\beta J(x)\tanh x. \tag{5.13.1}$$

It is obvious that the model exhibits a first order phase transition near $T = T_b$ (slightly below this temperate) (see Fig. 5.13.6); At $T_b$ a new non-zero fixed point appears. The stability of solutions may be read off from the bifurcation diagram Fig. 5.13.7.

Below $T_b$ there is a branch where $m$ is not zero. The possibility of hysteresis (e.g., supercooling) can also be found. An equilibrium phase transition (or the coexistence of two phases) occurs somewhere the branches corresponding to the coexisting phases are stable. To determine the exact phase transition point requires an analogue of

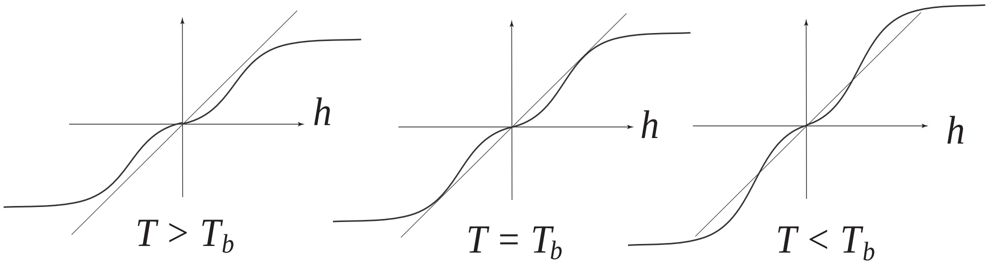


Fig. 5.13.6 A first order phase transition occurs slightly below $T_b$. To determine the exact phase transition temperature, we need an analogue of Maxwell's rule.

Maxwell's rule (see Fig. 4.2.1 or Fig. 4.2.2).

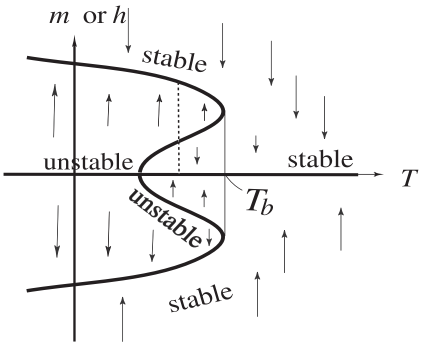


Fig. 5.13.7 The bifurcation diagram for the model that allows a first order phase transition. The vertical arrows denote the evolving direction of perturbation to the fixed point values of $m$ at various temperatures. We can at once see the stability of the fixed points from the exchange of stability occurring at every bifurcation. To determine the exact phase transition temperature (denoted by the dotted line in the figure; within the mean-field theory) we need a rule parallel to Maxwell's rule.

If the cooperation effect is overwhelmed by an external field, phase transition can occur. Below $T_c$ 2-Ising model is in the up phase or down phase. If a small magnetic field is applied, then the direction of the spins of one phase is stabilized relative to the other phase. Let us discuss the phase transition induced by this change with the aid of a mean field theory. Let us assume $J$ is constant as in Section 5.8, and we consider (5.8.7), i.e.,

$$x = 2d\beta J \tanh(x + \beta h). \tag{5.13.2}$$

Let us look at Fig. 5.13.8. Initially, there is no magnetic field, and the system is in the 'down' phase (spins are predominantly down; black disk). If a magnetic field pointing upward is applied, the true equilibrium state must be the 'up' phase. However, if the magnetic field is not too large, the island of up spins cannot stably exist in the ocean of down spins unless it is larger than some size (larger than the critical nucleus size), because the interface (phase boundary) free energy penalty is too large comparing to the decrease of free energy due to the favorable magnetic field. Thus, the down phase

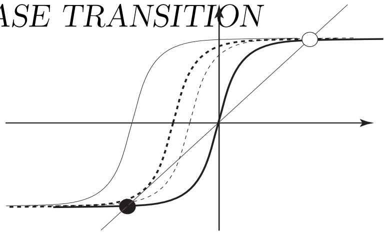


Fig. 5.13.8 If the magnetic filed intensity is increased (the curve moves to the left), the stability of the 'down' phase (black disk) diminishes, and becomes metastable until the curve reaches the thick dotted curve. If the magnetic field is increased further, this 'down' phase become unstable, and the phase jumps to the 'up' phase (white disk).

is *metastable*. If the magnetization becomes larger than the value corresponding to the thick dotted line in the figure the down phase becomes *unstable*, and without any nucleation the system goes to the up phase. If a bifurcation diagram is constructed with the magnetization $h$ as the parameter, it is easy to see what happens.

The picture just explained applies to many common first order phase transitions. Magnetic field is the conjugate field of magnetization. In parallel to the illustration in Fig. 5.3.3 for a fluid system (or a binary mixture system), pressure (or chemical potential) may be regarded as the conjugate variable to the order parameter. Thus, the first order phase transition that occurs when an external field is altered is very common.

To make microscopic order sensitive to fluctuations, one method is to increase the degree of freedom of the order parameter. From this point of view, an interesting generalization of the Ising model is the $p$-state *Potts model*, where $p$ is a positive integer. In this model, at each lattice point is a variable (Potts spin) that can take $p$ different states ($s_i \in \{1, \cdots, p\}$), and the system Hamiltonian reads

$$H = -J \sum_{\langle i,j \rangle} \delta_{s_i,s_j}, \tag{5.13.3}$$

where $J > 0$ is a coupling constant. In short, if two Potts spins on the nearest neighbor lattice points are in the same state, it is energetically stabilized; otherwise the pair is destabilized. The Ising model is a 2-state Potts model. If $p$ is large, then aligning spins should become difficult, so the second order phase transition we know for $p = 2$ would become first order for a large $p$. Indeed, this is true. In 2-space, it is known that the phase transition is first order for $p = 5, 6, \cdots$. In 3-space the transition is first order for $p = 3, 4, \cdots$.[112]

---

[112] A summary of mathematical results can be found in M. Biskup and L. Chase, Commun. Math. Phys., **238** 53 (2003).

In Fig. 4.2.2 (lower side) we studied the free energy of the van der Waals fluid in the metastable states. According to the figure, at least the free energy of the metastable states looks available by extrapolating the free energy of stable phases. This may give us a hope that the free energy of the metastable phase can be obtained by analytically continue (with respect to temperature) the stable phase free energy. It is known that this is impossible for short-range interaction systems.[113] However, if the interaction range is infinite as the Kac potential it is known that the above expectation is correct.[114] In the following, Andreev's argument for gas-liquid phase transition singularity is outlined.

The probability for a gas bubble of radius $R$ to be formed in the liquid phase is determined by the required minimum work to create the bubble:

$$\Delta W = (4\pi/3v)[\mu_g(T) - \mu_l(T)]R^3 + 4\pi\sigma R^2, \tag{5.13.4}$$

where $\sigma$ is the boundary free energy of the gas-liquid interface, and $v$ is the volume of gas per molecule. If $T > T_b$ (boiling point), the chemical potential $\mu_g$ of the gas phase is smaller than the chemical potential $\mu_l$ of the liquid phase, so for sufficiently large $R$

$$\frac{\partial \Delta W}{\partial R} < 0. \tag{5.13.5}$$

That is, once a large bubble is formed, it grows without limit. The number of bubbles of radius $R$ for $T < T_b$ is given by

$$n(R) \sim \exp[-\Delta W(R)/T]. \tag{5.13.6}$$

Each bubble of radius $R$ contributes the following free energy:

$$\varphi(R) = A(T_b - T)R^3 + BR^2, \tag{5.13.7}$$

where $A$ and $B$ are positive constants. Let $q$ be the evaporation heat. Note that

$$\mu_g(T) - \mu_l(T) = q(T_b - T)/T_b. \tag{5.13.8}$$

---

[113] S. Friedli and C.-E. Pfister, "On the Singularity of the Free Energy at a First Order Phase Transition," Commun. Math. Phys., **245**, 69 (2004). This was first discussed by A. F. Andreev, "Singularity of thermodynamic quantities at a first order phase transition point," Soviet Physics JETP **15**, 1415 (1964), and then proved by Isakov in 1984.

[114] S. Friedli and C.-E. Pfister, "Non-Analyticity and the van der Waals Limit," J. Stat. Phys., **114**, 665 (2004).

The contribution of all the bubbles to the free energy change is given by

$$\Delta\Phi = V\int_0^R 4\pi R^2 dR\varphi(R)n(R). \tag{5.13.9}$$

To estimate this when we approach the phase transition point from below (from the liquid side) we have only to evaluate the following integral:

$$\int_0^\infty R^m \exp\left[-A\frac{T_b - T}{T_b}R^3 - \frac{B}{T_b}R^2\right] dR. \tag{5.13.10}$$

This diverges for $T > T_b$. In the limit $T \to T_b - 0$ the derivatives of $\Delta\Phi$ with respect to $T$ converges for all orders. The formula (5.13.10) only allows us to estimate very high order derivatives (because, for small order derivatives the contribution of the integrand come from microscopic $R$), but for high order derivatives with the aid of an asymptotic evaluation of the $\Gamma$-function, we have

$$\frac{d^n\Phi(T_b-)}{dT^n} \simeq (-AT_b^{1/2}/B^{3/2})^n\sqrt{n}(3n/2e)^{3n/2}. \tag{5.13.11}$$

That is, $\Phi$ is not holomorphic, but $C^\infty$. If we introduce $\tau = i\eta$, we realize that (5.13.10) behaves as $\exp(c/\eta^2)$. In other words, the boiling point is a $C^\infty$ singularity, so generally speaking, there is no hope to obtain the free energy of the overheated liquid by analytical continuation.

## 5.14 Phenomenology of first order phase transition

The first order phase transition is phenomenologically discussed very often in terms of the Ginzburg-Landau Hamiltonian in physical chemistry. The basic idea and its relation to the standard statistical mechanics are reviewed to conclude this chapter.

For many first order phase transitions, fluctuations are not crucial, so it is natural to guess that mean-field theory is useful. As has already been stated, it is not easy to justify the results of mean field arguments, and it is not very wise to swallow the mean field results uncritically, but they are heuristically useful.

Let us write the (generalized) canonical partition function in terms of the Ginzburg-Landau Hamiltonian as

$$Z = \sum_{\phi} \exp\left\{-\int d^d\boldsymbol{r}\left[\frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}\tau_0\phi^2 + \frac{1}{4}g\phi^4 + h\phi\right]\right\}. \tag{5.14.1}$$

Here, the sum over $\phi$ is a formal representation of functional integral and means summation over all the order parameter field configurations. In the mean field approximation, fluctuations are ignored. In the bulk phase, $\phi$ must be spatially constant, so we ignore the derivative term. Therefore, the free energy density may be written as

$$\beta\mathcal{F}(\tau, h) = \frac{1}{2}\tau m^2 + \frac{1}{4}gm^4 + hm, \tag{5.14.2}$$

where $\phi$ is replaced by its spatially uniform value $m$ and the suffixes '0' have been removed. The basic idea is that the $m$ that minimizes $\mathcal{F}$ (perhaps with some other conditions) is realized in equilibrium.

Fig. 5.14.1 illustrates (5.14.2) as a function of $m$ for various values of $h$ and $\tau$.

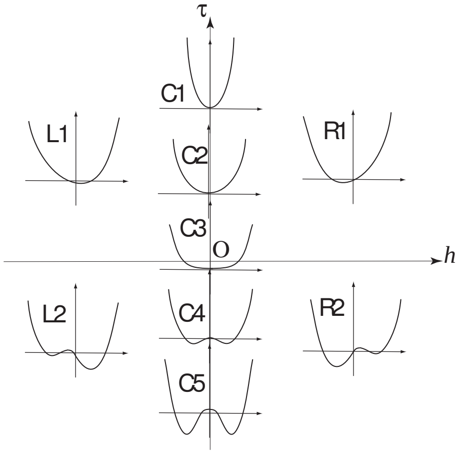


Fig. 5.14.1 Sketches of Landau free energy density (5.14.2).
For each free energy density graph the abscissa is the order parameter $m$. C3 is at the critical point. The local free energy has a flat bottom and fluctuation is large around $m = 0$. As temperature increases from C2 to C1, the disordered phase becomes clearer. In contrast, on the lower temperature side C4 and C5 two phases become possible. Whether one of them is realized or coexistence occurs depends on systems. As temperature is lowered, order phases become more ordered. L1 and R1 are under the bias of the conjugate variable in the high temperature phase. No qualitative change is seen. In contrast, in L2 and R2 phase transitions between ordered phases may be possible.

The reader must think immediately that the free energy function which is not convex as a function of the order parameter is nonsensical. This conclusion is correct. Therefore, let us call (5.14.2) the *quasi free energy density.* We have already encountered a similar situation in the case of the van der Waals fluid. The difficulty was evaded with the aid of Maxwell's rule there. A similar strategy may work in the general case, but here we take the meaning of the integrand of (5.14.1) seriously. It is the free energy at the mesoscopic scale (the free energy of a macroscopically small system; we could call it the *mesoscopic free energy*). In the preceding argument, we simply discarded the gradient term, and also discarded fluctuations as irrelevant. However, even without looking at Fig. 5.14.1 everybody knows that $m$ at the mesoscopic scale violently fluctuates, and cannot be a variable that specifies a thermodynamic state. This implies that to understand thermodynamics of the system we must sum up (5.14.1). However, what result we can obtain may be guessed fairly accurately and intuitively from the mean field results, that is, from Fig. 5.14.1.

To calculate the free energy as a function of the order parameter $m$, let us apply step by step coarse-graining just as in the case of real space renormalization, instead of trying to compute thermodynamic quantities at once. From this point of view, the quasi free energy above is the effective free energy we would obtain by coarse-graining fluctuations up to an intermediate mesoscopic scale. As is explained below, we can guess the free energy in the thermodynamic limit from this effective free energy, so quasi free energy is not a bad quantity to consider. However, to this end, we must first recognize that there are two kinds of order parameters, *non-conserved order parameters* (NCOP) and *conserved order parameters* (COP) .

For example, magnetization density of ferromagnets is a NCOP. If we cut out a finite spatial portion and isolate it, the sum of NCOP is not conserved; magnetization of the volume can change without any conservation constraint. In this case, we can simply minimize the obtained free energy to determine the phase diagram without paying any attention to any constraints. Therefore, even if there are two minima at some coarse-grained scale, if they have different free energy values, then further coarse-graining enlarges the discrepancy and the effective free energy makes a sharper and sharper valley around the smallest minimum, so no phase coexistence can occur in this case. In other words, the actual order parameter value observed in a macroscopic system is close to $m$ that gives the smallest minimum value in Fig. 5.14.1. Thus, we may say qualitative conclusions can be read off from Fig. 5.14.1. However, if the smallest minima are not unique, we cannot say what would happen by this approach. As in the case of sufficiently cold 3-Ising model there may be coexistence of multiple phases, or as the 2-Ising model, only one of them is realized without any coexistence.

A good example of the COP is the density in the case of gas-liquid phase transition and the concentration (difference) in the case of the binary fluid mixture. If we cut out a finite spatial portion and isolate it, the sum of COP is conserved; in these examples this is due to the conservation of materials. Since atoms and molecules cannot move at once to a distant place, the conservation low applies locally. To know what phases are realized in a given system, we must minimize the effective free energy appropriate to the scale under the constraint that the sum of $m$ over the system is given. Fig. 5.14.2 (upper side) illustrates how the effective free energy as a function of the order parameter evolves as the COP system is coarse-grained. The right most graphs correspond to the thermodynamic limit (in practice, the system size need be a few times as large as the correlation length; this is usually tiny). The flat parts we see here correspond to the flat portions in Fig. 5.13.1 and Fig. 5.3.2. The states marked with a and b correspond to thermodynamic pure phases (single phases), and the states between them correspond to coexisting phases. If we mimic this evolution, we can roughly guess how the free energy depends on the order parameter in the thermodynamic limit. Thus, it is usually not very difficult to guess the thermodynamic limit from the quasi free energy. Needless to say, however, the calculation of the actual value of phase transition temperature or order parameter values is out of question.

Suppose that the system-wise average of the order parameter density is $m$, and this value lies in the flat part of the free energy in the thermodynamic limit. What happens actually? Let us call the pure phase corresponding to 'a' in Fig. 5.14.2 as phase I, and 'b' as phase II. Let $\alpha$ be the fraction of phase I and the order parameter of phase I (resp., phase II) be $m_{\rm I}$ (resp., $m_{\rm II}$). $\alpha m_{\rm I} + (1-\alpha)m_{\rm II} = m$ must be kept constant. Under this condition the free energy density reads

$$\mathcal{F} = \alpha\mathcal{F}(m_{\rm I}) + (1-\alpha)\mathcal{F}(m_{\rm II}), \tag{5.14.3}$$

and $\alpha$ is determined geometrically by the so-called *lever rule* (Fig. 5.14.3). The pair $m_{\rm I}$ and $m_{\rm II}$ determines the phase coexistence curve on the phase diagram (cf. Fig. 5.14.7).

Is there any use of mesoscale effective free energies illustrated in Fig. 5.14.2 (lower left side or the starting points of the upper side) which is not convex? The effective free energy must tell us the fluctuation at the appropriate spatial scale. In the case of NCOP, we can trivially see what fluctuation we can observe. In the COP case, we must pay due attention to the conservation law. In this case, even for smaller scale fluctuations phase separation could be locally induced. As noted already, molecules cannot jump over a long distance, so they move only diffusively locally. Therefore, if the local average of the order parameter happens to be $m$, then this value cannot change rapidly (i.e., conservation law holds locally). Consequently, we have only to think what happens in a small volume (of the appropriate coarse-grained scale). If

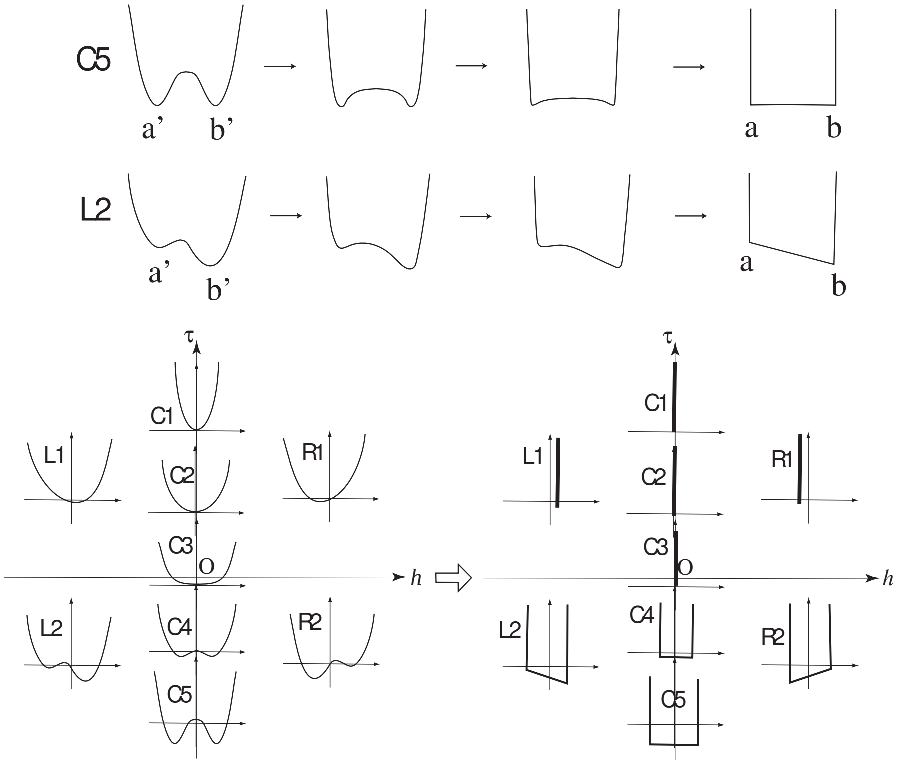


Fig. 5.14.2 How free energy changes as coarse-graining goes on. The upper figures illustrates for two examples how the effective free energy evolve as the system is coarse-grained. The leftmost figures are quasi free energy. The lower figure illustrates how the quasi free energies in Fig. 5.14.1 correspond to the true free energy (density) in the thermodynamic limit. For each graph the abscissa is the whole-system average value of $m$, and the ordinate denotes the free energy density. The state that cannot be realized in equilibrium is assigned $+\infty$; this convention is natural as can be guessed from the quasi free energy functions.

the effective free energy on this length scale is locally convex around the point P corresponding to the local average $m$, then the state is stable against small fluctuations. However, the total free energy might be reduced further by phase separation (see Fig. 5.14.4). If this is possible (i.e., the convex envelop of the effective free energy can have the portion below the effective free energy graph), the single phase corresponding to point P is said to be metastable. If the effective free energy is not convex locally around P, the single phase separates into two phases by small fluctuations,

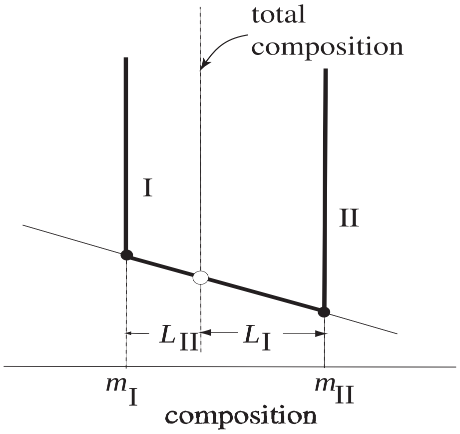


Fig. 5.14.3 The lever rule: The ordinate of this graph is the free energy (density) and the abscissa is the order parameter. Suppose the system-wise average of the order parameter is indicated by the dotted vertical line passing through the white disk. The ratio of phase I and phase II is given by the ratio of the lengths $L_{\rm I}$ and $L_{\rm II}$. $m_{\rm I}$ and $m_{\rm II}$ dictate the compositions of the coexisting phases. The ordinate of the white disk denote the free energy density.

so we say the local state is unstable.

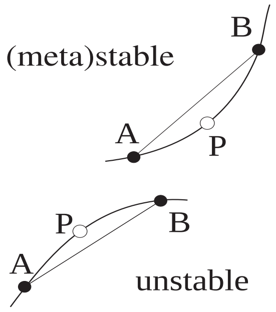


Fig. 5.14.4 If the effective free energy is locally convex around P, small fluctuation cannot segregate this into two phases, say at A and B, because such phase separation increases the effective free energy. Therefore, the state P is (meta)stable. Whether P is actually metastable or stable cannot be distinguished if we pay attention only to the neighborhood of P. If the effective free energy is not locally convex, the total free energy of the system can be reduced by a small fluctuation that separates P into A and B, so this state is unstable.

When phase separation occurs mesoscopically, the ratio of the separated phases can be determined by the *lever rule* just as before as explained in Fig. 5.14.5. In this figure the concentration difference is the order parameter, which is a COP.

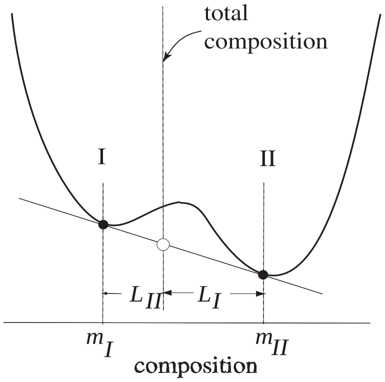


Fig. 5.14.5 The lever rule. Suppose the total composition of the mixture is given by the vertical broken line. The uniform state of this composition is unstable, so it separates into phase I and phase II to minimize the total free energy. The lowest total free energy realized is indicated by the open circle. The ratio of the phase I and II is given by the ratio of $L_I$ and $L_{II}$.

It must be emphasized that in the thermodynamic limit true equilibrium states of a system whose average $m$ is in the interval $(m_{\rm I}, m_{\rm II})$ is, as illustrated in Fig. 5.14.3 coexisting of phase I and II. However, even for a macroscopic system, since molecules cannot travel a long distance very quickly, if a mixture is metastable around the correlation length scale, truly equilibrium states are not so easily realized. Thus, understanding the stability of homogeneous state with various $m$ in terms of quasi free energy can be useful. Fig. 5.14.5 allows us us to guess the phase coexistence behavior away from the critical point as illustrated in Fig. 5.14.6. The result is the popular phase diagram Fig. 5.14.7.

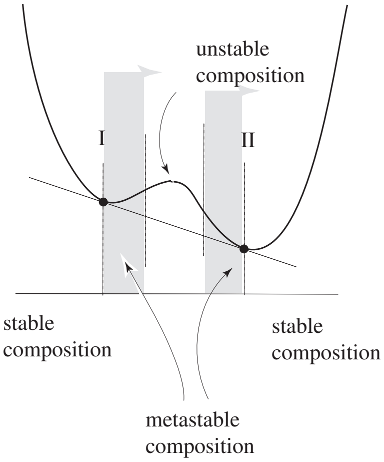


Fig. 5.14.6 Stable, metastable and unstable states for the COP case.
Where the convex envelop of the curve can come blow the original curve phase coexistence occurs. The fictitious single phase whose order parameter value $\langle m \rangle$ is the average value is metastable if the graph around $\langle m \rangle$ is convex. Otherwise, unstable.

The boundary between the metastable and unstable states is called the *spinodal curve*. This curve is practically definable if fluctuation is small, but it is not a well defined curve near the critical point.

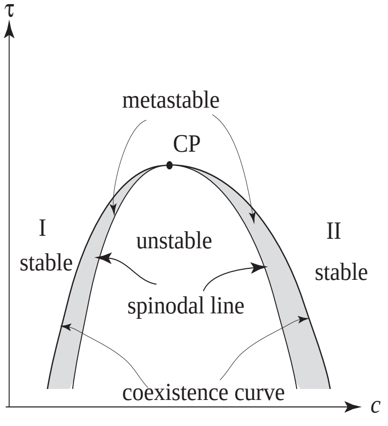


Fig. 5.14.7 The phase diagram of a binary mixture corresponding to Fig. 5.14.6. The gray portion is the metastable phase, and the white region below the spinodal curve is the unstable phase, but near the critical point CP the distinction is blurred due to large fluctuations.

**Supplementary Pages relevant to phase transitions**
At least the following topics are posted in the Supplementary Pages in a pedagogically accessible fashion.

a more systematic real space RG(not yet)
A concrete calculation illustration of field theoretical RG,
RG for Self-Avoiding walk (not yet),
2D Ising model exact solution
RG of the XY model (not yet)

## Exercises for Chapter 5

**5.1** [Phase transition and analyticity]
If there is no phase transition in a range of fugacity $z = e^{\beta\mu}$, $P/k_BT$ is, by definition, holomorphic in $z$, so we may expand it as

$$\frac{P}{k_BT} = \sum_{\ell=1}^{\infty} b_\ell z^\ell, \tag{5.P.1}$$

where $b_\ell$ is called the $\ell$-body cluster coefficient. They are smooth (actually real analytic) functions of $T$ and positive for lower temperatures, so each $b_\ell$ has the smallest real positive zero $T_\ell$. It is known that $\{T_\ell\}$ is a monotone decreasing sequence of $\ell$. It is demonstrated[115] that
(i) $b_\ell(T_c) > 0$ if $\ell$ is sufficiently large.
(ii) There are infinitely many $T_\ell$ between any $T$ $(> T_c)$ and $T_c$.
Let $T_a$ be the accumulation point of $\{T_\ell\}$. Show $T_c = T_a$.[116] **r**

**5.2** [Crude version of rigorous Peierls' argument] **r**
Let us impose an all up spin boundary condition to the 2-Ising model on the finite square. Then, we wish to take a thermodynamic limit. If the spin at the center of the square is more likely to be up than to be down, we may conclude that there is a long-range order.

Let $\gamma$ be a closed Bloch wall (i.e., the boundary between up and down spin domains; this does not mean that the domain enclosed by $\gamma$ is a totally up or down domain (lakes can contain islands with ponds with islets, etc.; the wall corresponds to the shore lines.) The probability $P_V(\gamma)$ to find such a wall in the system with volume $V$ has the following estimate (we used this in our discussion on Peierls' argument):

$$P_V(\gamma) \le e^{-2\beta J|\gamma|},$$

where $|\gamma|$ is the total length of the contour $\gamma$, $\beta = 1/k_BT$, and $J$ is the usual ferromagnetic coupling constant. [This naturally looking inequality needs a proof; it is not trivial.]
(1) Since the outside boundary is all up, there must be a Bloch wall encircling the origin for the spin at the origin to be down. Therefore, the probability $P_V^0$ of the spin at the origin to be down must be smaller than the probability of the occurrence of at least one contour encircling the origin. Show

$$P_V^0 \le \sum_{\gamma} e^{-2\beta J|\gamma|}, \tag{5.P.2}$$

---

[115] T. Kihara and J. Okutani, Chem. Phys. Lett., **8**, 63 (1971).

[116] This problem asks a mathematically trivial question, but the main point is the fact stated here.

where the summation is over all the possible contours surrounding the origin. [Hint: Don't think too much. If event $A$ occurs when at least one of $B$ and $C$ occurs, then $A \subset B \cup C$.]
(2) Estimate the number of contours with circumference $|\gamma|$ crudely as the total number of random walks of length $|\gamma|$ starting from appropriate neighborhood points of the origin. Use this crude estimate and show that if $\beta$ is sufficiently large, $P_V^0 < 1/2$ for large $V$.

**5.3** [Phase transition in 1D long-range system]
Using Peierls' argument, discuss the phase transition in a 1d spin system whose coupling constant behaves as $r^{-q}$ ($q < 2$) beyond some distance $r_0$ (you may assume that the coupling constant for $r < r_0$ is $J$, constant). No rigorous argument is wanted.

**5.4** [Griffiths' inequality]
Empirically, it is known that there is the following relation among critical exponents:

$$\alpha + \beta(1 + \delta) = 2. \tag{5.P.3}$$

(1) Thermodynamically, demonstrate the following inequality (*Griffiths' inequality*)

$$\alpha + \beta(1 + \delta) \geq 2. \tag{5.P.4}$$

[Hint: You may proceed just as the case of Rushbrooke's inequality, but use $m \sim h^{1/\delta}$ (at $T = T_c$) to differentiate wrt $H$ under constant temperature. At $\tau = 0$ and $h = 0$ the (vertical) line $T = T_c$ is tangent to $m = m(\tau, h = 0)$, so $m$ may be parameterized by $\tau$.]
(2) Using the scaling relation or Kadanoff's argument (that is, using the expression of the magnetization as a generalized homogeneous function of $h$ and $\tau$), show that the equality actually holds. [Hint: the wisest approach may be to use $\beta\delta = \beta + 1$. You may use such as $\alpha + 2\beta + \gamma = 2$ we have already discussed.]

**5.5** [A toy illustration of Lee-Yang theory][117]
Suppose the grand partition function of a fluid in a volume $V$ is given by

$$\Xi_V = (1 + z)^V \frac{1 - z^V}{1 - z}, \tag{5.P.5}$$

where $z$ is the fugacity.
(1) Find the zeros of $\Xi_V$. How does the distribution of the zeros change as $V \to \infty$?
(2) Obtain $P$ for real positive $z$ and locate the phase transition.
(3) Find the volume $v$ per particle as a function of $z$.

---

[117] S similar question can be found in Reichl

(4) Find the $P$-$v$ relation, and demonstrate that phase coexistence does happen.

**5.6** [The Lee-Yang circle theorem illustrated
The the theorem is proved in Supplementary Pages. Here, let us check its content for simple cases.
(1) According to the theorem the root of $p(z) = 1 + 2az + z^2$ must be on the unit circle as long as $a \in [-1, 1]$. Confirm this.
(2) On the apices of a triangle are spins interacting with each other. For this system construct the polynomial of $z$, and confirm that as long as the interactions are ferromagnetic, all the zeros are on the unit circle.

**5.7** [A derivation of mean field theory]
A mean field approach may be obtained with the aid of a variational principle for free energy. If the (density) distribution function of microstates is $f$ (we consider classical case) the Helmholtz free energy may be written as:

$$A = \langle H \rangle + k_B T \int d\Gamma f(\Gamma) \log f(\Gamma). \tag{5.P.6}$$

Here, the integration is over the whole phase space. $\langle H \rangle$ is the expectation value of the system Hamiltonian with respect to $f$. Let us apply this to the Ising model on a $N \times N$ square lattice. Its Hamiltonian is as usual

$$H = -J \sum_{\langle i,j \rangle} S_i S_j. \tag{5.P.7}$$

If we could vary $f$ unconditionally and minimize $A$, then the minimum must be the correct free energy, but this is in many cases extremely hard or plainly impossible. Therefore, we assume an approximate form for $f$ and the range of variation is narrowed. For example, we could introduce a 'single-body' approximation:

$$f = \phi(S_1)\phi(S_2) \cdots \phi(S_N), \tag{5.P.8}$$

where $\phi$ is a single-spin (density) distribution function.
(1) Under this approximation write down $A$ in terms of $\phi$. That is, find $X_1$ and $X_2$ in the following formula:

$$A = X_1 \left[ \sum_S \phi(S) S \right]^2 + X_2 \sum_S \phi(S) \log \phi(S), \tag{5.P.9}$$

where $N$ and $N \pm 1$ need not be distinguished.
(2) Minimize $A$ wrt $\phi$. $\phi$ must be normalized. What is the equation determining $\phi$?

(3) Using the obtained formula, write down the magnetization per spin. The used Lagrange's multiplier must be determined.

**5.8** [Gibbs-Bogoliubov' inequality and mean field]
(1) Derive the following inequality (called the Gibbs-Bogoliubov inequality) with the aid of Jensen's inequality for classical systems [This question has already been asked in Chapter2, but is reproduced here]:

$$A \leq A_0 + \langle H - H_0 \rangle_0. \tag{5.P.10}$$

Here, $A$ is the free energy of the system with the Hamiltonian $H$, $A_0$ is the free energy of the system with the Hamiltonian $H_0$, $\langle\ \rangle_0$ is the average over the canonical distribution wrt $H_0$. [Hint: compute $\left\langle e^{-(H-H_0)}\right\rangle_0$; the temperature may be absorbed or we could use the unit system with $k_B = 1$.]

All the variational approximations for statistical thermodynamics are applications of this inequality.[118] Let $H$ be the Hamiltonian of the system we are interested in, and $H_0$ be the Hamiltonian of a system whose free energy $A_0$ we can compute exactly. We introduce variational parameters in $H_0$ and tehn try to make the RHS of (6.5.43) as small as possible.
(2) As $H$ we adopt the $N \times N$ 2-Ising model Hamiltonian (without a magnetic field; even with it there is almost no change), and

$$H_0 = \sum_i h s_i. \tag{5.P.11}$$

Derive the equation for $h$ that minimizes the RHS of (6.5.43).

**5.9** [Exact mean field for 1-Ising model]
The starting point of the mean-field theory can be the following exact relation for the 1-Ising model:

$$\langle s_0 \rangle = \langle \tanh \beta J(s_{-1} + s_1) \rangle, \tag{5.P.12}$$

where $\langle\ \rangle$ is the equilibrium expectation. Utilizing $s^2 = 1$ and translational symmetry of the system, write down a closed equation for $m = \langle s \rangle$, and then discuss the possible phase transitions.

**5.10** [2-Ising model on the honeycomb lattice]
Let us consider a 2-Ising model on the honeycomb lattice whose coupling constant is $J$. Assume there is no magnetic field.
(1) Find the equation corresponding to (5.8.4).

---

[118] See, for example, M. D. Girardeau and R. M. Mazur, "Variational methods in statistical mechanics," Adv. Chem. Phys. XXIV, eds. I. Prigogine and S.A. Rice (Academic, New York, 1974), p187-255.

(2) Find $T_c$ with the aid of the approximation corresponding to (5.8.6).
(3) Then, using a more accurate mean field theory corresponding to (5.8.13) compute $T_c$. Which $T_c$ obtained by (2) or this question should be lower? Is your result consistent with your expectation?

**5.11** [1-Gaussian model]
At each lattice point $i$ of a one-dimensional lattice lives a real variable $q_i$, and the system Hamiltonian is given by

$$H = \sum_j \left[\frac{1}{2}q_j^2 - Kq_jq_{j+1}\right]. \tag{5.P.13}$$

The partition function reads

$$Z = \left(\prod_j \int_{-\infty}^{\infty} dq_j\right) \prod_{j=1}^{N-1} \exp[w(q_j, q_{j+1})], \tag{5.P.14}$$

where

$$w(x, y) = -\frac{1}{4}(x^2 + y^2) + Kxy. \tag{5.P.15}$$

The partition function should be evaluated just as the 1-Ising model with the aid of the eigenvalue problem:

$$\lambda f(x) = \int_{-\infty}^{\infty} dy\, f(y) \exp\left[-\frac{1}{4}(x^2 + y^2) + Kxy\right]. \tag{5.P.16}$$

The integral kernel is Gaussian, so the eigenfunction belonging to the largest eigenvalue should be of constant sign [Hint: you can see a correspondence to the transfer matrix approach; actually, there is a counterpart of Perron-Frobenius theorem for positive definite integral kernels]. Therefore we may assume that $f$ is also Gaussian.
(1) Find $f(x)$ (its multiplicative numerical coefficient may be ignored).
(2) Find the free energy per lattice point. Is thee any phase transition?

**5.12** [Correlation function by mean field theory, or mean field for nonuniform space]
(1) Let us assume that the coupling constant and the magnetic field depend on spatial position:

$$H = -\sum_{\langle i,j\rangle} J_{ij}s_is_j - \sum_i h_is_i. \tag{5.P.17}$$

Derive the basic equation for the mean-field theory for a square lattice:

$$\langle s_i\rangle = \left\langle \tanh(\beta \sum_j J_{ij}s_j + \beta h_i)\right\rangle. \tag{5.P.18}$$

If we introduce the crude approximation like (5.8.6), we obtain

$$\langle s_i \rangle = \tanh(\beta \sum_j J_{ij} \langle s_j \rangle + \beta h_i). \tag{5.P.19}$$

(2) We wish to compute the spatial correlation $\langle s_i s_j \rangle$. First, demonstrate that

$$\frac{\partial \langle s_i \rangle}{\partial h_k} = k_B T \langle s_i s_k \rangle \tag{5.P.20}$$

without any approximation.
(3) Applying this to the following form of (6.5.79), obtain the equation for $\{\langle s_i s_k \rangle\}$:

$$\mathrm{Arctanh} \langle s_i \rangle = \beta \sum_j J_{ij} \langle s_j \rangle + \beta h_i. \tag{5.P.21}$$

(4) Now, let us go over to the continuum limit, assuming that the system has a translational symmetry. If we write the correlation as $g$, the equation obtained in (3) becomes

$$\int dy \left( \frac{\delta(x-y)}{1-m^2} - \beta J(x-y) \right) g(y-z) = \delta(x-z). \tag{5.P.22}$$

We have already assumed the spatial translational symmetry and set $m(x) = m$. The Fourier transform of the coupling constant reads

$$J(k) = \sum_j e^{ik \cdot r_j} J(r_j) = 2^d J \cos k_x \cos k_y \cdots . \tag{5.P.23}$$

Find the Fourier transform $G(k)$ of the correlation function $g$.
(5) If we are interested in global features, we have only to pay attention to small $k$. Determine the coefficients $A$ and $B$ in the following formula (you may assume $T > T_c$):

$$G(k) \simeq \frac{1}{A + Bk^2}. \tag{5.P.24}$$

(6) Determine the critical exponent $\nu$.

**5.13** [Lattice gas on honeycomb lattice]
Let us relate the 2-Ising model on the honeycomb lattice and the lattice gas on the

same lattice. The Ising Hamiltonian $H$ and the lattice gas Hamiltonian $H_L$ as just as given in the text:

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_i s_i, \tag{5.P.25}$$

$$H_L = -J' \sum_{\langle i,j \rangle} n_i n_j. \tag{5.P.26}$$

Let $V$ ($\gg 1$) be the total number of lattice points and down spins are regarded particles.

(1) Following the procedure around p276, rewrite the canonical partition function of the Ising model in therms of the number of down spins $[D]$ and that of down spin pairs $[DD]$.

(2) Express the lattice gas pressure in terms of magnetic field $h$ and the free energy per spin $f$.

(3) Demonstrate that the lattice gas pressure $P$ is a continuous function of $h$.

(4) Sketch the free energy $Vf = -k_BT \log Z$[119] of the Ising model for a few representative temperatures. Next, sketch the pressure of the lattice gas as a function of $\log z$ (this is essentially the chemical potential) for a few representative temperatures. Then, explain their noteworthy features succinctly.

**5.14** [RG by Migdal approximation[120]]

When we discussed 'decimation', we have realized that the procedure is not very good in the space higher than 1D. For example, if we apply the method to the 2-Ising model (taking $\ell = 2$, i.e., thin half of spins), we obtain

$$K' = \frac{1}{4} \log \cosh 4K, \tag{5.P.27}$$

where the Hamiltonian is written in the following form:

$$H = -\sum K s_i s_j \tag{5.P.28}$$

and the temperature is absorbed in the parameter. The fixed point of this transformation is $K = 0$ (i.e., the high temperature limit), so there is no ordering.

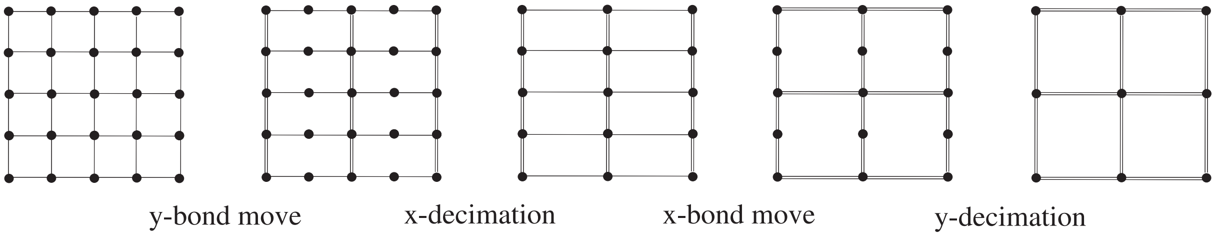
Migdal proposed to remedy the defect of underestimating the interactions as follows (see the figure below).

(i) [$y$-bond moving step] Every other vertical bonds ($y$-bonds) are combined with their right-neighboring bonds . If the coupling constant in the $y$-direction is $K_y$, the

---

[119] Notice that this free energy is $G$ rather than $A$, since $h$ is on.

[120]



coupling constant due to the new bonds made by combining two vertical couplings is $2K_y$.
(ii) [$x$-decimation step] For the $x$-direction, one dimensional thinning is performed. For the new $x$-directional coupling constant is computed by the 1D thinning result we obtained (5.10.10).
(iii) [$x$-bond moving step] Next, every other $x$-bonds are merged with their lower neighbor $x$-bonds.
(iv) [$y$-decimation step] Apply one-dimensional decimation in the $y$-direction.
Thus, we have arrived at the square lattice with the lattice spacings doubled (i.e., $\ell = 2$). If we halve the spatial scale we can complete a renormalization group transformation.
(1) Let us put $'$ to the parameters after the procedure (i)-(iv). Show that

$$K_x' = \log\cosh(2K_x), \tag{5.P.29}$$

$$K_y' = \frac{1}{2}\log\cosh(4K_y). \tag{5.P.30}$$

Here, the 'initial values' are $K$ for both the $x$ and $y$ couplings. Notice that in two different directions, the step-cycle of the procedure is 'out of phase,' so to speak. In the $x$-direction, the decimation is applied first and then the bond are moved, while in the $y$-direction the bonds are merged first, and then decimation follows. Consequently, the fixed points of these two equations have different fixed points (marked with $^*$): $K_x^* = 2K_y^*$.
(2) Find all the fixed points $K_x^*$. Which corresponds to the critical fixed point?
(3) Linearizing the renormalization transformation around the fixed point, we can calculate critical exponents; we have only to compute $d\log\cosh(2K_x)/dK_x$. This corresponds to $\ell^{y_1}$. Determine $\nu$. We can not say the result is impressive, but still there is an improvement from the mean-field approach. **5.8** ['Democracy'][121]
Assume that there is a hierarchical organization whose decision at the level $k$ is determined by $s$ units at the level $k-1$. Let us assume that level 0 corresponds

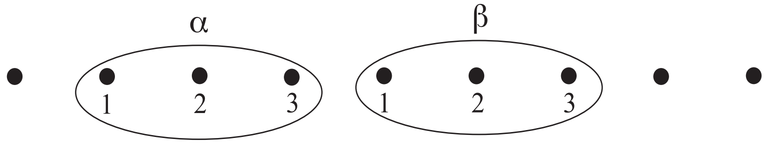
---

[121] cf. S. Galam and S. Wonczk, "Dictatorship from majority rule voting," Eur. Phys. J. B **18**, 183 (2000).

to individual persons, and level $F$ is the decision of the organization; we assume there are sufficiently many levels (in practice, 5 is enough). This is sort of a coarse-graining system of people's opinions; any political system is a coarse-graining system of opinions, so no subtle voice of conscience can go through to the top.
(1) Suppose $s = 3$ and the strict majority rule is used. Find all the fixed points of this system, and tell which is stable and which is not.
(2) If $s = 4$ under the majority rule but when opinion A ties with B, always A is chosen. What is the minimum fraction of the people with opinion B to dominate the organization? How about the $s = 2$ case?
(3) Suppose $s = 5$ with strict majority rule, but there is an organizational propaganda for one opinion. Because of this, there is definitely one member in $s$ at each level that its opinion is fixed. What is the minimum percentage of the anti-organizational opinion to prevail?

**5.15** [Finite size effect]
The specific heat of a certain magnetic system behaves $C \sim |\tau|^{-\alpha}$ near its critical point without external magnetic field, if the specimen is sufficiently large. If the magnet is not very large, or more concretely, if it is a sphere of radius $R$, near its critical point, its maximum specific heat is $C(R)$. Compute the ratio $C(2R)/C(R)$ in terms of critical indices.

**5.16** [Use of block spins in 1-Ising model]
Let us construct an RG transformation for a 1-Ising model with a similar approach as is applied to the triangle lattice1-Ising model. We start with (5.10.12). The figure corresponding to Fig. 5.10.2 os as shown below:



The equation corresponding to (5.10.13) is

$$K's'_{\alpha}s'_{\beta} = Ks_{\alpha 3}s_{\beta 1}, \tag{5.P.31}$$

and $s'$ is the block spin of ±1 determined by the majority rule. This relation cannot literally be realized, so just as in the triangle lattice case, we need an analogue of (5.10.14).
(1) Write down $\phi(K)$ corresponding to (5.10.16).
(2) Write down the RG equation corresponding to (5.10.19) and (5.10.21).
(3) Find the fixed points.

(4) What can you conclude from these calculations?

**5.17** [‘Democracy’][122]
Let us consider a hierarchical organization in which the decision at the $k$th level depends on the decisions of the $s$ cells of the $k-1$st level. Assume that the 0th level corresponds to individual members of the organization, and the decision at level $F$ is the decision of the organization. If there are sufficiently many levels (actually 5 levels are enough), the system may be understood as a system to coarse-grain individual opinions. To be frank, any political organization is a coarse-graining mechanism of opinions, and it is usually the case that conscientious subtle voices do not reach the top.

In the following we assume there are two options A and B that must be chosen. Consider the fraction $p_n$ of the cells at level $n$ that support A.

(1) Suppose $s = 3$ and strict majority rule is applied. Find all the fixed points of this system and study their stability.

(2) Suppose $s = 4$. Majority rule is applied but if two opinions A and B are equally supported, A is always selected. For B to be the decision of the organization, at least how many % of the people should support B? In the extreme case, if $s = 2$ what happens?

(3) Suppose $s = 5$. Majority rule is applied, but due to the organizational propaganda at every level there is always at least one cell that supports A. For B to win despite this arrangement, what is the minimum % of the supporters of B?

---

[122] cf. S. Galam and S. Wonczk, “Dictatorship from majority rule voting,” Eur. Phys. J. B **18**, 183 (2000). The paper contains some trivial calculation errors, so trust your own result.

# Chapter 6

# Solutions

In this solution set, to come back from the destination of hyperlinks click q.

## 6.1 Problems for Chapter 1

**1.1** [Equivalence of heat and work]
A block of mass $M = 1$ g is at rest in space (vacuum). Another block of the same mass and velocity $V = 1.5$ km/s collides with the first block and the two blocks stick to each other.
(1) Assuming that there is no radiation loss of energy and that there is no rotation of the resultant single block, obtain the temperature of the resultant single block after equilibration. Assume that the specific heat of the material is 2.1 J/g·K.
(2) If rotation of the resultant body is allowed, what can be said about its final temperature? In particular, is it possible not to raise the temperature of the resultant single block? (Only a qualitative discussion will be enough.)
**Solution**
(1) The total initial macroscopic kinetic energy is $MV^2/2 = 1.125 \times 10^3$ J. The final total kinetic energy of macroscopic motion is (the necessary speed is determined by

the conservation of linear momentum)

$$\frac{1}{2}(2M)(V/2)^2 = \frac{1}{4}MV^2 = 562.5\text{J}. \tag{6.1.1}$$

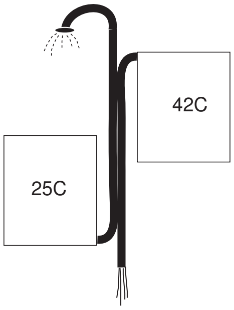
Therefore, 562.5 J should become the energy of thermal motion. Thus, 562.5/4.2 = 134 K is the increase in temperature, so the final temperature is 334 K.
(2) Rotational motion can be excited, so the temperature increase is reduced. However, this rotation is due to the non-zero angular momentum around the center of mass of the initial system. Now, the question is whether the rotational kinetic energy can preserve the kinetic energy of relative motion. If the second body has an extremely long thin rod to connect it to the other body to become a single block, then we can reduce the loss of rotational kinetic energy indefinitely (compute the final rotational kinetic energy and compare it with the relative kinetic energy). That is, the temperature increase can be made indefinitely small.

**1.2** [Exchange of temperatures]
Suppose there are two water tanks A and B containing the same amount of water. Initially, A is 42°C and B is 25°C. The final state we want is A to be 25°C and B 42°C (that is, the temperatures of A and B are exchanged; e.g., A is the used warm bath water, and B new clean tap water). Assume that the whole system is thermally isolated.
(1) Is the process reversibly feasible? Explain why.
(2) Devise a process. No precise statement is needed; only state key ideas.
(3) If the reader's process in (2) contains mechanical parts (e.g., a heat engine), devise a process without any moving parts. No precise statement is needed; only state key ideas. The reader can use other containers to hold water and can freely move or flow water (ignore dissipation).
**Solution**
(1) The initial and final states obviously have the same entropy. Thus, if there is a way to connect these two in a quasiequilibrium fashion, it is reversible. Well, is there any quasiequilibrium process connecting them? [Notice that thermodynamics, esp., the second law postulates that there is an adiabatic process connecting two equilibrium states of the same system; the process must inevitably reversible in this case. This is usually an assumption; however, depending on the formalism of equilibrium thermodynamics, this can be proved.[1]]
(2) Operate a reversible engine between the two tanks until the temperatures become equal. The work produced may be stored by pulling up a weight. Now, use

---

[1] E. H. Lieb and J. Yngvason, "The physics and mathematics of the second law of thermodynamics," Phys. Rep., **310**, 1-96 (1999). This is the state of the art of thermodynamics.

the engine as a heat pump with the aid of the stored energy in the weight.
(3) The answer (2) is fine, but if you wish to commercialize the machinery, you need an engine and a motor, not very economical. In any case moving parts are where troubles start (as is often the case with the hard disk). Thus, we wish to get rid of moving parts. In the actual commercial product, water from the faucet is guided through some clever heat exchange device. Thus, we assume we can freely move water (but slowly). We must reduce the production of entropy, so heat transfer between different temperatures should be maximally avoided. Initially, the temperatures are distinct, so this is impossible. However, we can make this initial mismatch effect indefinitely small by making the amount of water to be equal temperature from both tanks as small as possible. Look at the scheme in the following figure. If the pipe is thin enough, the effect of initial awkward thermal contact is reduced as much as you wish, and the subsequent heat contact can be as isothermal as possible.[2]



The thermal contact between counter flows can be made indefinitely long by making a tortuous route. Cooled water is drained and warmed water is used for your shower. Such a device is industrially common to save energy,

**1.3** [The fourth law of thermodynamics]
(1) For 0.5 moles of a certain substance the equation of state is obtained as:

$$E = \kappa T V^{1/2}, \tag{6.1.2}$$

where $E$ is internal energy, $V$ is volume, $T$ is absolute temperature, and $\kappa$ is a constant. Write down the equation of state for $N$ moles of this substance.
(2) We can define extensive quantities per molecule $x = X/N$, where $X = E, S, V$ and $x = e, s, v$. Write down the Gibbs relation for one mole (or a molecule) That is, express $de$ in terms of $x$ and other extensive quantities per mole (or molecule).

**Solution**
(1) Let $e$ and $v$ be internal energy per mole and the volume per mole. The given equation of state can be rewritten as

$$e/2 = \kappa T(v/2)^{1/2} \Rightarrow e = \sqrt{2}\kappa T v^{1/2}, \tag{6.1.3}$$

---

[2] There is a way to do gradual temperature change through preparing numerous heat baths with various temperatures, but this is of course only for extremely rich people.

so

$$(E/N) = \sqrt{2}\kappa T(V/N)^{1/2} \Rightarrow E = \sqrt{2}\kappa T N^{1/2}V^{1/2}. \tag{6.1.4}$$

(2) You could use $d(E/N) = dE/N - (E/N)(dN/N)$ and $E = TS - PV + \mu N$ (because $E$ is extensive), but a wiser method is to use

$$e = sT - Pv + \mu \tag{6.1.5}$$

and the Gibbs-Duhem relation, $SdT - VdP + Nd\mu = 0$ or

$$sdT - vdP + d\mu = 0. \tag{6.1.6}$$

Differentiating (6.1.5) and using (6.1.6), we get

$$de = Tds - Pdv. \tag{6.1.7}$$

**1.4** [Asymmetric coin and limit theorem]
The law of large numbers applies to an unfair coin as well. Suppose the coin we use has a probability $1/2 + \varepsilon$ to yield a head (H). Otherwise, a tail (T) is yielded. One get \$1 when H shows up and must pay \$1, otherwise.
(1) Write down the generating function $\omega_N(k)$ for the reader's expected gain per one coin-tossing $s_N$ for the length $N$ coin-tossing sequence.
(2) Compute the mean and the variance of $s_N$.
(3) Using the generating function technique, find the density distribution function $f_N$ for the fluctuation of $s_N$ to demonstrate[3] the law of large numbers and the central limit theorem.

**Solution**
(1) Let $X_i$ be your gain by the $i$-th tossing. Then, $s_N = (1/N)\sum X_i$.

$$\begin{aligned}
\omega_N(k) &\equiv \left\langle \exp\left(ik\frac{1}{N}\sum X_i\right)\right\rangle, && (6.1.8)\\
&= \prod^N \left\langle \exp\left(ik\frac{1}{N}X_i\right)\right\rangle, \text{ due to statistical independence} && (6.1.9)\\
&= \omega(k/N)^N, && (6.1.10)
\end{aligned}$$

where $\omega$ is the generating function for the single tossing:

$$\begin{aligned}
\omega(k) &= \left(\frac{1}{2}+\varepsilon\right)e^{ik} + \left(\frac{1}{2}-\varepsilon\right)e^{-ik}, && (6.1.11)\\
&= \cos k + 2i\varepsilon \sin k. && (6.1.12)
\end{aligned}$$

---

[3] Need not be mathematical; quite a theoretical physicist's way is OK!

(There is no point to streamline this expression.) Therefore,

$$\omega_N(k) = [\cos(k/N) + 2i\varepsilon \sin(k/N)]^N. \tag{6.1.13}$$

(2) [Directly] This is straightforward: the mean is

$$M_N = \langle s_N \rangle = \frac{1}{N}\sum \langle X_i \rangle = (1/2+\varepsilon) - (1/2-\varepsilon) = 2\varepsilon. \tag{6.1.14}$$

The variance is:

$$\begin{aligned} V_N = \langle (s_N - \langle s_N \rangle)^2 \rangle &= \left\langle \left( \frac{1}{N}\sum (X_i - 2\varepsilon) \right)^2 \right\rangle, & (6.1.15) \\ &= \frac{1}{N^2}\left\langle \sum_i (X_i - 2\varepsilon)^2 + 2\sum_{i>j}(X_i - 2\varepsilon)(X_j - 2\varepsilon) \right\rangle. & (6.1.16) \end{aligned}$$

$X_i$ and $X_j$ are statistically independent, so we can average them separately. Therefore, the cross terms in the above vanish and we obtain

$$V_N = \frac{1}{N}\langle (X_1 - 2\varepsilon)^2 \rangle = \frac{1}{N}(\langle X_1^2 \rangle - 4\varepsilon^2) = \frac{1}{N}(1 - 4\varepsilon^2). \tag{6.1.17}$$

[Using the generating function] We use the generating function: (note that $\omega(0) = 1$)

$$M_N = \left. \frac{d\omega_N(k)}{dik} \right|_{k=0} = -iN[\cdots]^{N-1}\left[ -\frac{1}{N}\sin\frac{k}{N} + 2i\varepsilon\frac{1}{N}\cos\frac{k}{N} \right]_{k=0} = 2\varepsilon. \tag{6.1.18}$$

The variance is obtained by the logarithmic second derivative (the derivative is evaluated at $k = 0$ at the end; the calculation may be easier if you do it in terms of $e^{\pm ik}$):

$$\begin{aligned} V_N &= -\frac{d^2}{dk^2}\log \omega_N(k) = -N\frac{d^2}{dk^2}\log[\cos(k/N) + 2i\varepsilon \sin(k/N)], & (6.1.19) \\ &= -\frac{d}{dk}\frac{-\sin(k/N) + 2i\varepsilon\cos(k/N)}{\cos(k/N) + 2i\varepsilon\sin(k/N)}, & (6.1.20) \\ &= \frac{1}{N}\left\{ \frac{\cos(k/N) + 2i\varepsilon\sin(k/N)}{\cos(k/N) + 2i\varepsilon\sin(k/N)} + \frac{[-\sin(k/N) + 2i\varepsilon\cos(k/N)]^2}{[\cos(k/N) + 2i\varepsilon\sin(k/N)]^2} \right\}_{k=0}, & (6.1.21) \\ &= \frac{1}{N}(1 - 4\varepsilon^2). & (6.1.22) \end{aligned}$$

Here, in (6.1.21) the sign in front of the second term is +, because we have an overall − before differentiation.
(3) Assuming $N$ is large, so $k/N$ is small, we Taylor expand $\omega_N(k)$. It is far cleverer to expand $\log \omega$ (partition function!; you should see how partition functions are mathematically natural objects), because we have computed the derivatives.

$$\log \omega_N(k) = ikM_N - \frac{1}{2}k^2 V_N + o[k^2]. \tag{6.1.23}$$

Therefore, the inverse-Fourier transformation gives

$$\begin{aligned} f_N(x) &= \frac{1}{2\pi i}\int dk e^{-ikx}\omega_N(k) = \frac{1}{2\pi i}\int dk \exp\left(-ikx + ikM_N - (k^2/2)V_N\right), & (6.1.24)\\ &= \frac{1}{2\pi i}\int dk \exp\left(-(V_N/2)[k^2 - 2ik(x-M_N)/V_N]\right), & (6.1.25)\\ &= \frac{1}{2\pi i}\int dk \exp\left(-(V_N/2)[k - i(x-M_N)/V_N]^2 - (1/2V_N)(x-M_N)^2\right), & (6.1.26)\\ &\propto \exp\left(-(1/2V_N)(x-M_N)^2\right). & (6.1.27) \end{aligned}$$

Needless to say, this is consistent with our calculation above, but this tells us that the distribution is Gaussian with variance $V_N = O[1/N]$; **the central limit theorem is a refinement of the law of large numbers**.

**1.5** [How to use Chebyshev's inequality]
(1) We wish to know whether a coin is fair or not. To estimate the probability of H within ±0.01, how many throwings do you need? Let us tolerate larger errors once in 100 runs. You may assume that the coin is not grossly unfair.
(2) Theoretically, it is known that if the coin is flipped rapidly, the final surface agrees with the original surface with probability 0.51 (for example, if the original surface is H, then with probability about 0.51 H is obtained). To confirm this bias how many throwings do you think is needed?

**Solution**
(1) Let $p$ be the true probability for H which is not very different from 1/2. Let $\chi_i$ be the index function of the $i$-th trial to be up. Then

$$p = E(\chi_1).$$

Chebyshev's inequality tells us

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\chi_i - p\right| > \varepsilon\right) < \frac{1}{n\varepsilon^2}V(\chi_1).$$

We know $V(\chi_1) = p - p^2 = p(1-p) \simeq 1/4$. Now, $\varepsilon = 0.01$ and the overall error tolerance is 0.01. Therefore, we may assume

$$\frac{1}{n\varepsilon^2}V(\chi_1) = \frac{10^4}{4n} = 0.01.$$

Therefore, $n = 2.5 \times 10^5$ is needed. That is, if if you throw the coin 250,000 times, you can estimate the head probability of the coin within the tolerance about ±2%. Or, with confidence level 99% you can find $p$ within 2%.
(2) The answer depends on the level of your precision demand. Since the unfairness is of 1% order, you must be able to estimate $p$ at least to this order. The answer to (1) gives a practical answer, if you reduce the confidence level to, say, 95%. If you wish to stick to the 99% confidence level, then 250,000 trials are not enough; $\varepsilon$ should be halved at least, and you need $10^6$ trials. See Diaconis et al., "Dynamical Bias in the Coin Toss," SIAM Review **49**, 211 (2007). The lesson is, "If we have this much trouble analyzing a common coin toss, the reader can imagine the difficulty we have with interpreting typical stochastic assumptions in an econometric analysis."

**1.6** [A Monte Carlo method to determine $\pi$]
There is a unit square ($1 \times 1$ square) and in it is inscribed a disk of radius 1/2 sharing the centers of mass. Randomly dropping points uniformly on the square, and counting the number of points landing on the disk, we can measure $\pi$ (or rather $\pi/4$ directly). How many points do we need to obtain 3 digits below the decimal point of $\pi/4$ with probability more than 99%?

**Solution**
Let us introduce an iid variable $X$ such that $X = 1$ if the point lands on the disk and 0, otherwise, and uniformly distributed on the unit square. Then, we expect

$$\langle X \rangle = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} X_i = \pi/4. \tag{6.1.28}$$

The variance of $X$ is $\pi(1-\pi/4)/4 < 3/16$. Therefore, Chebyshev tells us that

$$Prob.\left(\left|\frac{1}{N}\sum X_i - \frac{\pi}{4}\right| > 10^{-3}\right) \le \frac{3/16}{10^{-6}N}. \tag{6.1.29}$$

This upperbound should be less than $10^{-2}$, so $N \simeq 2 \times 10^7$ is needed.

**1.7** [Law of large numbers does not hold, if the distribution is too broad (if fluctuations are too wild)]

The Cauchy distribution that is defined by the following density distribution function

$$p(x) = \frac{1}{\pi}\frac{a}{x^2+a^2}. \tag{6.1.30}$$

does not satisfy $E(|X|) < +\infty$ (needless to say, the variance is infinite). Actually, the density distribution of

$$E_n = \frac{X_1 + \cdots + X_n}{n} \tag{6.1.31}$$

has exactly the same distribution function as $X_1$, if $\{X_j\}$ all obey the same Cauchy distribution and are statistically independent. Let us demonstrate this.
(1) What is the characteristic function of the Cauchy distribution? You can look up the result, but even in that case you must explain why the result is correct.
(2) Show what we wish to demonstrate.

**Solution**
(1)
We have only to compute

$$\omega(k) = \int_{-\infty}^{\infty} dx\, e^{ikx} p(x) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} dx \left(\frac{1}{x-ia} - \frac{1}{x+ia}\right) e^{ikx} = e^{-a|k|}. \tag{6.1.32}$$

It may be a good occasion to review contour integration, Cauchy's theorem, etc.
(2) The characteristic function for $E_n$ is given by $\omega(k/n)^n$. This is in our case exactly $\omega(k)$ itself. QED!

**1.8** [St. Petersburg Paradox by Daniel Bernoulli]
Let $\{X_i\}$ be iid with

$$P(X_1 = 2^n) = 2^{-n} \tag{6.1.33}$$

for all positive integers $n$.
(1) Show that $E(X_1) = \infty$.
Thus, it seems that if $X_1$ is the gambler's gain, the reader can participate in this gambling game with any entry price and still can expect a positive gain. However, any sensible person would pay \$1000 as a fair price for playing. Why? This is the 'paradox.'
(2) Needless to say, the law of large numbers does not hold for $E_n$. This implies that empirically obtainable expectation and theoretical one should have some discrepancy. Indeed, it can be proved (the reader need not show this; not very easy) that for any positive $\varepsilon$

$$P\left(|E_n/\log_2 n - 1| > \varepsilon\right) \to 0 \tag{6.1.34}$$

in the $n \to \infty$ limit. Recall that $E_n$ is the expected payoff. Explain why the reader does not with to pay \$1000. (Or for this to be a fair price how many times does the reader have to play?)

**Solution**

(1) This is obvious:

$$E(X_1) = \sum_m 2^m 2^{-m}. \tag{6.1.35}$$

(2) The above estimate implies with high probability (asymptotically) $E_n \sim \log_2 n$. That is, you must wait until $\log_2 n = 1000$ to have break-even. That is $n = 2^{1000} \simeq 10^{300}$.

**1.9**. Explain 'Bertrand's paradox' in about 10 lines (without using outrageous fonts). What lesson can you learn? [You can easily find a detailed account in the web.]

**Solution**

Wikipedia: `http://en.wikipedia.org/wiki/Bertrand's_paradox_(probability)` gives a good account of this topic. In short, being random or uniform sampling is a rather tricky idea. We need a detailed empirical analysis of what we mean by 'uniform' or 'random'. This is the lesson.

However, this article may have given you an idea that there is a general principle to 'rescue' the ambiguity inherent in the concept of 'lack of knowledge' following Jaynes ('maximum ignorance principle' or, in essence, to use fully the symmetry in the problem). This implies that we must perform a detailed analysis of what is NOT known. If symmetry principles are used inadvertently, we can easily get nonsensical result. A classic example is *von Mises' wine/water paradox*. You can look this up in the web, and perhaps will see proposals to resolve the paradox. The resolutions require more detailed analysis of what is not known.

**1.10** [System with dissipation]

There is a classical particle system described by the canonical coordinates $\{q, p\}$ ($q$ and $p$ are collective notations for position and momentum coordinates of all the particles in the system). In terms of the Poisson bracket and the system Hamiltonian the equation of motion may be written as

$$\frac{dq}{dt} = [q, H], \quad \frac{dp}{dt} = [p, H] - \alpha p,$$

where $\alpha$ is a small positive constant. That is, the system is not perfectly classical mechanical, but contains a certain effect of dissipation.[4]

(1) Demonstrate that the Liouville's theorem is violated for this system.

(2) Demonstrate that the system energy decreases. (Assume that $H = K + U$ as

---

[4] This model may look artificial, but similar models are used to study nonequilibrium systems.

usual and $K$ is quadratic in $p$.)
(3) However, show that if $H$ is time dependent, then the system energy may be maintained.

**Solution**
(1) Let us check the incompressibility of the flow defined by this differential equation:

$$\frac{\partial}{\partial q}\frac{dq}{dt}+\frac{\partial}{\partial p}\frac{dp}{dt}=-\alpha\frac{\partial p}{\partial p}=-3N\alpha, \tag{6.1.36}$$

where $N$ is the number of the point particles and 3 is the spatial dimensionality ($3N$ is the number of $p$ coordinates).
(2) Let us compute

$$\frac{dH}{dt}=\frac{\partial H}{\partial t}+[H,H]+\frac{\partial H}{\partial p}\cdot(-\alpha p)=-2\alpha K, \tag{6.1.37}$$

where $K$ is the kinetic energy and is positive definite. Therefore, $H$ decreases. Here $K$ being a quadratic form (a homogeneous function of order 2) is used.
(3) As you can see from (6.1.37), if $H$ is explicitly time dependent, you can do whatever you want. (For example, although it is not physically interesting, $H(t)=U+e^{2\alpha t}K$ would do.)

**1.11** The following can be read in a textbook. Point out the error in the argument.
“In general, there is no logical “room” for adding extra assumptions, such as equal a priori probability. The evolution of an actual system is determined by the laws of mechanics (or quantum mechanics). If the results of using any extra assumptions always agree with the logical consequence of the laws of mechanics, and it should be possible to show that fact. If they do not agree with the laws of mechanics, then the extra assumptions are wrong.”

**Solution**
Dynamical laws are differential laws, so they must not be discussed without auxiliary conditions such as initial conditions. Such conditions are given independent of the dynamical laws. That is, there is a room to introduce extra assumptions. For example, when we sample from the phase volume, it seems that we can only sample according to he probability measure absolutely continuous with respect to the Riemann volume, but this has nothing to do with the dynamical laws. However, when we study chaotic dynamical systems, observability condition of chaos is a condition on the initial condition.

As can be seen from this example, it is simple minded or too haste to conclude that some universal properties applicable to systems obeying dynamical laws is only due to dynamical laws. Another example is the unavoidable external disturbances. Therefore, small stochastic perturbations are always there. That this is the reason for the statistical nature of the system is logically perfectly legitimate assertion. If the disturbance is universal, then the results would also be universal.

To be more precise, there is no verification of dynamical laws for many body systems. Therefore, at least purely logically no one can conclude that "If they do not agree with the laws of mechanics, then the extra assumptions are wrong."

**1.12** Classically, the microcanonical distribution may be written as

$$\hat{W}(E) = \frac{1}{N!h^{3N}} \int \delta(H(q,p) - E) dq dp. \tag{6.1.38}$$

Show that this can be expressed as follows:

$$\hat{W}(E) = \frac{1}{N!h^{3N}|grad\, H|} d\sigma, \tag{6.1.39}$$

where $\sigma$ is the area element of the constant energy surface and $grad\, H$ is the $3N$-dimensional gradient vector of the system Hamiltonian $H$ with respect to the canonical variables $(q_i, p_i)$.

**Solution**

Notice that $dpdq = d\sigma dE$. If the $6N - 1$ dimensional canonical coordinates in the surface is collectively written as $\boldsymbol{\sigma}$, then $H$ can be written in terms of $E$ and $\boldsymbol{\sigma}$. Notice that the following formula for the $d$-function holds:

$$\delta(f(E) - E_0) = \frac{1}{|f'(E_0)|} \delta(E - E_0). \tag{6.1.40}$$

If as $f(E)$ $H(E, \boldsymbol{\sigma})$ is adopted, then we must use the partial derivative of $H$ with respect to $E$ as $f'$, but it is the directional derivative $H$ perpendicular to the isoenergetic surface, so its absolute value must be identical to the absolute value of the gradient. Thus we have demonstrated (6.1.39).

**1.13** [Equipartition of energy with the aid of microcanonical ensemble]

Within the microcanonical ensemble formalism[5] for a classical fluid consisting of $N$ interacting but identical particles,

(1) Show that the temperature $T$ can be expressed in terms of the average of the reciprocal kinetic energy as ($N \gg 1$ may be assumed)

$$k_B T = \frac{2}{3N\langle K^{-1} \rangle}, \tag{6.1.41}$$

---

[5] Use $W(E)$, the phase volume of the states with the total energy not exceeding $E$.

where $K$ is the total kinetic energy of the system.
Comment: We are NOT discussing ideal systems. The system can have any interaction among particles. $T$ is defined thermodynamically as $1/T = \partial S/\partial E$.
(2) In the thermodynamic limit show that this is equivalent to

$$k_B T = \frac{2}{3}\langle \kappa \rangle, \tag{6.1.42}$$

where $\kappa$ is the average kinetic energy par particle. This is essentially the equipartition of kinetic energy. [Hint. the reader may use intuitively the weak law of large numbers.]

**Solution**
(1) Although I asked the case with $N \gg 1$ to use $W(E)$ instead of the energy shell volume (i.e., $W(E + \delta E) - W(E)$), if you carefully do a similar calculation below, you get the result true for small $N$ as well.

The phase volume $W(E)$ of the states $H = \sum p_i^2/2m + V(q) \le E$ is computed as

$$W(E) = \int_{V(q)\le E} dq \int_{\sum p_i^2/2m \le E - V(q)} dp = \int_{V(q)\le E} dq\, C_N \left[m(E - V(q))\right]^{3N/2}, \tag{6.1.43}$$

where $C_N$ is a geometrical factor we need not calculate. Temperature is defined by

$$\frac{1}{T} = k_B \frac{d}{dE}\log W(E) = \frac{3N}{2}\frac{\int dq\,(E - V(q))^{3N/2-1}}{\int dq\,(E - V(q))^{3N/2}}. \tag{6.1.44}$$

You may have expected the derivative to consist of two terms, but the derivative with respect to the integration range for $\{q\}$ vanishes, because the integrand vanishes there: $V(q) = E$. The above formula is the average of $(E - V(q))^{-1}$ over the configuration space. It is not yet the desired result that is an average over the phase space. Now, we use (again) the fact that for high-dimensions the total volume is almost on the skin. Therefore,

$$\int_{K\le\mathcal{E}} dp\, K^{-1} = \frac{1}{\mathcal{E}}\int_{K\le\mathcal{E}} dp. \tag{6.1.45}$$

That is, the configuration space average and the phase space average agree:[6]

$$\beta = \frac{3}{2}N\langle K^{-1}\rangle. \tag{6.1.46}$$

---

[6] If you use the energy shell, perhaps the first step in the above calculation may be slightly more complicated, but this last step is trivial.

(2) Let $k_i$ be the kinetic energy of the $i$-th particle. $K = \sum k_i$. Notice that the law of large numbers implies

$$N\langle K^{-1}\rangle = N\left\langle \frac{1}{\sum k_i}\right\rangle = \left\langle \frac{1}{\sum_i k_i/N}\right\rangle \to \frac{1}{\langle k_i\rangle} = \frac{1}{\langle K\rangle/N} = N/\langle K\rangle. \tag{6.1.47}$$

Here, we intuitively assumed that $k_i$ are iid random variables.

**1.14** [Generalized homogeneous function]
As we will learn much later, various thermodynamic quantities diverge near the second-order phase transition point (critical point). The free energy density $f$ as a function of temperature $\tau \propto T - T_c$ and pressure $p$ behaves as[7]

$$f(\lambda^{y_1}\tau, \lambda^{y_2}p) = \lambda^d f(\tau, p), \tag{6.1.48}$$

where $\lambda$ ($> 0$) is a scaling factor, $d$ is the spatial dimensionality, and $y_1$ and $y_2$ are real constants (related to the so-called critical exponents). That is, $f$ is a generalized homogeneous function.[8]

Formulate the counterpart of Euler's theorem and demonstrate it. You may freely use the method of characteristics.

**Solution**
(1) Differentiating the given formula wrt $\lambda$ and subsequently setting $\lambda = 1$ gives the quasilinear PDE:

$$y_1\tau\frac{\partial f}{\partial \tau} + y_2 p\frac{\partial f}{\partial p} = d\,f. \tag{6.1.49}$$

This must be a necessary and sufficient condition for a differentiable function $f$ to satisfy the (generalized) homogeneity relation. To prove sufficiency, we must solve (6.1.49) with the aid of the method of characteristics.[9] The characteristic equation reads

$$\frac{d\tau}{y_1\tau} = \frac{dp}{y_2 p} = \frac{df}{d\,f}. \tag{6.1.50}$$

You may combine these three fractions in any way, but here let us choose the least sophisticated combinations:

$$\frac{d\tau}{y_1\tau} = \frac{dp}{y_2 p},\ \frac{d\tau}{y_1\tau} = \frac{df}{d\,f}. \tag{6.1.51}$$

---

[7] Precisely speaking, this is the singular part of the free energy as we will learn later. Peculiar phenomena near the critical point are governed by this part of the free energy.

[8] B. Widom realized from the empirical data that if $f$ is a generalized homogeneous function, then critical peculiar phenomena can be explained in a unified fashion.

[9] See Supplements

The general solution to these ODEs are

$$\tau^{1/y_1}/p^{1/y_2} = C_1,\ \ \tau^{d/y_1}/f = C_2, \tag{6.1.52}$$

where $C_1$ and $C_2$ are integration constants. Hence, the general solution to (6.1.50) is given by

$$f(\tau, h) = \tau^{d/y_1} g(\tau^{1/y_1}/p^{1/y_2}), \tag{6.1.53}$$

where $g$ is a well-behaved function (differentiable, bounded, etc.). Indeed,

$$f(\lambda^{y_1}\tau, \lambda^{y_2}p) = (\lambda^{y_1}\tau)^{d/y_1} g((\lambda^{y_1}\tau)^{1/y_1}/(\lambda^{y_2}p)^{1/y_2}) = \lambda^d \tau^{d/y_1} g(\lambda\tau^{1/y_1}/\lambda p^{1/y_2}) = \lambda^d f(\tau, p). \tag{6.1.54}$$

Another (much smarter) way to solve the characteristic equation (6.1.50) is to introduced the so-called 'dilation parameter' $\lambda$ as follows:

$$\frac{d\tau}{y_1\tau} = \frac{dp}{y_2 p} = \frac{df}{d\,f} = \frac{d\lambda}{\lambda}. \tag{6.1.55}$$

This is solved as the following three ODEs:

$$\frac{d\tau}{y_1\tau} = \frac{d\lambda}{\lambda},\ = \frac{dp}{y_2 p} = \frac{d\lambda}{\lambda},\ \ \frac{df}{d\,f} = \frac{d\lambda}{\lambda}. \tag{6.1.56}$$

The general solutions are

$$\tau/\lambda^{y_1} = C_1,\ \ p/\lambda^{y_2} = C_2,\ \ f/\lambda^d = C_3, \tag{6.1.57}$$

where $C_i$ are integration constants. Thus, the general solution to (6.1.49) reads

$$F(\tau/\lambda^{y_1}, p/\lambda^{y_2}, f/\lambda^d) = 0, \tag{6.1.58}$$

where $F$ is a well-behaved function. In other words, for any $\lambda$ ($> 0$)

$$f = \lambda^d G(\tau/\lambda^{y_1}, p/\lambda^{y_2}). \tag{6.1.59}$$

Here, $G$ is an appropriate differentiable function.

**1.15**. [Mixing entropy and Gibbs' paradox]
We have two ideal gases with the same volume $V$, pressure $P$, and temperature $T$. These two gases consist of different chemical species. Assume the whole system is thermally isolated during the following processes.
(1) Two boxes containing the above gases are connected. That is, now the total volume of the mixture is $2V$. Find the entropy change due to this procedure of joining

two boxes.
(2) Find the entropy change if two gases are mixed into a single volume of $V$.
(3) How can you actually measure the entropy change in (1) experimentally?

**Solution**

(1) See the illustration at the end.
Before answering the question, let us consider the indistinguishable case. The entropy formula for the gas in box $V$ is

$$S_1 = Nk_B\left(\log\frac{V}{N} + \frac{3}{2}\log\frac{E}{N} + c\right). \tag{6.1.60}$$

If two boxes are joined, then

$$S_{1+2} = 2Nk_B\left(\log\frac{2V}{2N} + \frac{3}{2}\log\frac{2E}{2N} + c\right) = 2S_1 \tag{6.1.61}$$

as expected.

Now, the distinguishable case.

$$S_{1+2} = (2N)k_B\left(\log\frac{2V}{N} + \frac{3}{2}\log\frac{E}{N} + c\right) = 2S_1 + 2Nk_B\log 2 \tag{6.1.62}$$

This can be considered as the superposition of each gas expanded to $2V$:

$$S_1' = Nk_B\left(\log\frac{2V}{N} + \frac{3}{2}\log\frac{E}{N} + c\right) = S_1 + Nk_B\log 2. \tag{6.1.63}$$

Superposing these two, we have

$$S_{1+2} = 2S_1 + 2Nk_B\log 2. \tag{6.1.64}$$

(2) Indistinguishable case:
Obviously

$$S_{1+2} = 2Nk_B\left(\log\frac{V}{2N} + \frac{3}{2}\log\frac{2E}{2N} + c\right) = 2S_1 - 2Nk_B\log 2. \tag{6.1.65}$$

This is the entropy decrease due to compression = halving the volume.

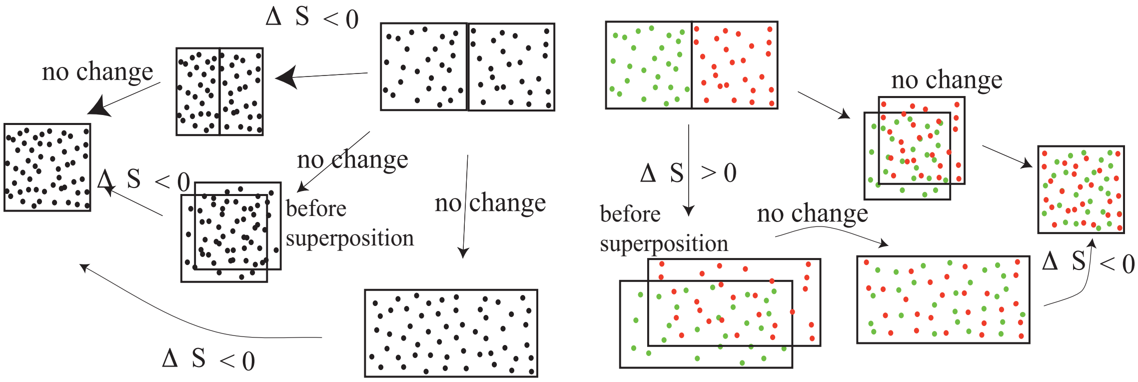
The distinguishable case: This is a simple superposition, so

$$S_{1+2}' = 2S_1 \tag{6.1.66}$$

That is, this is the result of Problem (1) compressed to the half volume: the entropy decrease due to compression is $2Nk_B \log 2$, which exactly cancels the mixing entropy.
(3) Notice that there is no exchange of heat during the mixing process in either of the two cases, distinguishable or indistinguishable. Therefore, to measure the mixing entropy you must somehow connect the mixed final state to the initial separated state (in a quasistatic way) and study this de-mixing process: for example, how much work do you have to supply to do this (or gain by doing this). There is a possibility of obtaining the absolute entropy of a gas with the aid of the Sackur-Tetrode formula. However, there is no way to use the result for the present purpose.

The above conclusion implies that to measure the mixing entropy you must be able to separate the mixture. If you did not know it is indeed a mixture, there would be no way to separate the 'mixture.' That is, you must be able to distinguish the components to measure the mixing entropy. You cannot use this experiment to tell whether two gases are identical or not.

You might say someone gave you a semipermeable membrane to separate the two components. However, since you have the membrane you can distinguish the components already.



**1.16** [To check that Boltzmann's formula does not contradict thermodynamics]
Let us check that Boltzmann's principle (within classical physics) is indeed consistent with thermodynamics: that is, if $S = k_B \log w(E, V)$,

$$dS = \frac{dE + PdV}{T}, \tag{6.1.67}$$

where $w(E, V)$ is the number of microstates satisfying that the energy is in $(E - \delta E, E]$ and the volume is in $(V - \delta V, V]$. Here, we clearly know what $E$ and $V$ are in both mechanics and in thermodynamics. The pressure $P$ can be computed mechanically, and $T$ is related to the average kinetic energy $K$ of the system.

Using the Boltzmann formula, we can write

$$dS = k_B \frac{1}{w}\frac{\partial w}{\partial E}dE + k_B \frac{1}{w}\frac{\partial w}{\partial V}dV. \tag{6.1.68}$$

Therefore, if we can compute partial derivatives in the above and identify their meanings, we should accomplish what we desire. This is actually what Boltzmann did in 1864. The demonstration is not very trivial, so here we wish to use the following relation

$$k_B \frac{1}{w}\frac{\partial w(E,V)}{\partial V} \to \frac{P}{T} \tag{6.1.69}$$

(in the thermodynamic limit) and consider only the energy derivative. We can write

$$w(E,V) = \int_{[E]} d\boldsymbol{q}d\boldsymbol{p} - \int_{[E-\delta E]} d\boldsymbol{q}d\boldsymbol{p} = \delta E \frac{\partial}{\partial E}\int_{[E]} d\boldsymbol{q}d\boldsymbol{p}, \tag{6.1.70}$$

where $[E]$ denotes the phase volume with energy not larger than $E$. We assume that the gas is confined in the volume $V$. Let $E = K(\boldsymbol{p}) + U(\boldsymbol{q})$, where $K$ is the total kinetic energy and$U$ the total intermolecular potential energy. The phase integration may be written as

$$\int_{[E]} d\boldsymbol{q}d\boldsymbol{p} = \int d\boldsymbol{q} \int_{K(\boldsymbol{p}) \le E - U(\boldsymbol{q})} d\boldsymbol{p}. \tag{6.1.71}$$

Thus, the integration with respect to $\boldsymbol{p}$ is the calculation of the volume of the $3N$-sphere of radius $\sqrt{2m(E-U(\boldsymbol{q}))}$.

(1) Show that

$$\frac{\partial}{\partial E}\int_{[E]} d\boldsymbol{q}d\boldsymbol{p} = \int d\boldsymbol{q}\frac{S_{3N}}{3N}2m\frac{3N}{2}[2m(E-U(\boldsymbol{q}))]^{3N/2-1}, \tag{6.1.72}$$

where $S_{3N}$ is the surface area of the $3N-1$-dimensional unit sphere.

(2) Using this formula, we can differentiate the integrand with $E$. Obtain

$$\frac{1}{w(E,V)}\frac{\partial w(E,V)}{\partial E} = k_B\left(\frac{3N}{2} - 1\right)\left\langle \frac{1}{K(\boldsymbol{p})}\right\rangle. \tag{6.1.73}$$

(3) We know from the kinetic theory that the average kinetic energy of a point particle is proportional to $T$ (precisely speaking, the average of $\boldsymbol{p}^2/2m = 3k_BT/2$). Assuming that all the kinetic energies of the particles are statistically independent,[10]

---

[10] This is not really a trivial statement; we need that the system is 'normal.' That is, the intermolecular interaction range must be very short, and the interactions are sufficiently repulsive in the very short range.

demonstrate that the formula in (2) is indeed equal to $1/T$.

**Solution**

(1) The integration with respect to $\boldsymbol{p}$ in (6.1.71) is the calculation of the volume of the $3N$-sphere of radius $\sqrt{2m(E-U(q))}$.

$$\int_{[E]} d\boldsymbol{q}d\boldsymbol{p} = \int d\boldsymbol{q} S_{3N} \int_0^{\sqrt{2m(E-U(q))}} p^{3N-1}dp = \int d\boldsymbol{q}\frac{S_{3N}}{3N}[2m(E-U(q))]^{3N/2}, \tag{6.1.74}$$

where $S_{3N}$ is the area of the unit $3N-1$-sphere. From this,

$$\frac{\partial}{\partial E}\int_{[E]} d\boldsymbol{q}d\boldsymbol{p} = \int d\boldsymbol{q}\frac{S_{3N}}{3N}2m\frac{3N}{2}[2m(E-U(q))]^{3N/2-1} \tag{6.1.75}$$

(2) The result of (1) gives us $w$, so we must differentiate this once more.

$$\begin{aligned}
\frac{\partial^2}{\partial E^2}\int_{[E]} d\boldsymbol{q}d\boldsymbol{p} &= \int d\boldsymbol{q}\frac{S_{3N}}{3N}(2m)^2\frac{3N}{2}\left(\frac{3N}{2}-1\right)[2m(E-U(q))]^{3N/2-2}, && (6.1.76)\\
&= \int d\boldsymbol{q}\frac{S_{3N}}{3N}(2m)\frac{3N}{2}\left(\frac{3N}{2}-1\right)[2m(E-U(q))]^{3N/2-1}\frac{1}{K(p)}, && (6.1.77)\\
&= \frac{\partial}{\partial E}\int d\boldsymbol{q}\frac{S_{3N}}{3N}\left(\frac{3N}{2}-1\right)[2m(E-U(q))]^{3N/2}\frac{1}{K(p)}, && (6.1.78)\\
&= \left(\frac{3N}{2}-1\right)w(E,V)\left\langle\frac{1}{K(p)}\right\rangle. && (6.1.79)
\end{aligned}$$

That is,

$$\frac{1}{w(E,V)}\frac{\partial w(E,V)}{\partial E} = \left(\frac{3N}{2}-1\right)\left\langle\frac{1}{K(p)}\right\rangle. \tag{6.1.80}$$

(3) We wish to demonstrate in the $N\to\infty$ limit

$$\left(\frac{3N}{2}-1\right)\left\langle\frac{1}{K(p)}\right\rangle = \frac{1}{k_BT}. \tag{6.1.81}$$

Obviously,

$$\left(\frac{3N}{2}-1\right)\left\langle\frac{1}{K(p)}\right\rangle = \left\langle\frac{1}{(3K(p)/2)/N}\right\rangle. \tag{6.1.82}$$

The law of large numbers tells us (cf. **1.13**) that

$$3K(p)/2)/N = \frac{3}{2}\frac{1}{N}\sum\frac{p^2}{2m} \to \frac{3}{2}\left\langle\frac{p^2}{2m}\right\rangle = k_BT. \tag{6.1.83}$$

**1.17** [Superrelativistic ideal gas]
Consider a super-relativistic gas consisting of particles whose energy $\varepsilon = c|\boldsymbol{p}|$, where $c$ is the speed of light, and $\boldsymbol{p}$ is the particle translational momentum.
(1) We have learned that the equation of state and the constant volume specific heat $C_V$ of an ideal gas may be obtained almost dimensional analytically, if we accept the basic postulate of statistical mechanics. Following this logic, find the pressure and $C_V$.
(2) Calculate the entropy to determine the constant corresponding to '$c$' (the constant in the entropy formula).

**Solution**
(1) Classically, we have only to consider the phase volume $\tilde{W}(V,E)$ whose dimension is $[q]^{3N}[p]^{3N}$. Therefore,

$$\tilde{W}(V,E) \sim V^N E^{3N}. \tag{6.1.84}$$

(Quantum mechanically, $[W] = 1 = [q]^{3N}/[h/p]^{3N} \sim [V]^N[E]^{3N}$.) This implies that

$$S = k_B \log \tilde{W}(V,E) = Nk_B(\log V + 3\log E + c_N) = Nk_B\left[\log\frac{V}{N} + 3\log\frac{E}{N} + c\right]. \tag{6.1.85}$$

Therefore,

$$\frac{P}{T} = \left.\frac{\partial S}{\partial V}\right|_E = \frac{Nk_B}{V} \Rightarrow PV = Nk_BT, \tag{6.1.86}$$

and

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_V = \frac{3NkB}{E} \Rightarrow E = 3NkBT. \tag{6.1.87}$$

Therefore, $C_V = 3Nk_B$.

(2) Let us do this honestly. $|\boldsymbol{p}| = h|\boldsymbol{n}|/2L$, where $L$ is the edge length of the box: $L^3 = V$, and $\boldsymbol{n}$ is a 3-dim positive integer vector. Therefore,

$$\tilde{w}(E,V) = \frac{1}{N!}\sum_{\{\boldsymbol{n}_i\}:\sum|\boldsymbol{n}_i|\simeq 2LE/ch} 1. \tag{6.1.88}$$

Here, $\simeq$ implies the value between $E-\delta E$ and $E$ (thin shell); in this calculation this is easier). This is not easy to evaluate directly, but a standard trick is to use the

Laplace transformation (or to use the generating function):[11]

$$\omega(t) = \int_0^\infty dE\, \tilde{w}(E,V)e^{-tE}. \tag{6.1.89}$$

Using (6.1.88), we can compute this as

$$\omega(t) = \frac{1}{N!}\int_0^\infty dE \sum_{\{\boldsymbol{n}_i\}:\sum|\boldsymbol{n}_i|\simeq 2LE/ch} e^{-\sum|\boldsymbol{n}_i|cht/2L} = \frac{1}{N!}\left(\sum_{\boldsymbol{n}} e^{-|\boldsymbol{n}_i|cht/2L}\right)^N. \tag{6.1.90}$$

The sum may be calculated as an integral over the first octant in the 3D space:

$$\sum_{\boldsymbol{n}} e^{-|\boldsymbol{n}_i|cht/2L} = \frac{1}{8}\int_0^\infty dn\, 4\pi n^2 e^{-|\boldsymbol{n}_i|cht/2L} = (8\pi)\left(\frac{L}{cht}\right)^3 \tag{6.1.91}$$

Therefore,

$$\omega(t) = \frac{1}{N!}(8\pi)^N\left(\frac{L}{cht}\right)^{3N}. \tag{6.1.92}$$

This implies (inverse Laplace transform; see an appropriate table. We ignore the difference between $nN$ and $nN-1$)

$$\tilde{w}(E,V) = \left(\frac{8\pi}{c^3h^3}\right)^N \frac{V^N}{N!}\frac{E^{3N}}{(3N)!}. \tag{6.1.93}$$

With the aid of Stirling's formula this reads

$$\tilde{w}(E,V) = \left(\frac{eV}{N}\right)^N \left(\frac{eE}{3N}\right)^{3N} \left(\frac{8\pi}{c^3h^3}\right)^N. \tag{6.1.94}$$

Therefore, entropy reads

$$S = Nk_B\left(\log\frac{V}{N} + 3\log\frac{E}{N} + \log\frac{8e^4\pi}{9c^3h^3}\right). \tag{6.1.95}$$

---

[11] Laplace transformation is not so popular among physicists, but it is only a variant of Fourier transformation. However, it is extremely useful in solving linear constant coef ODEs. See Applicable Analysis
`http://web.me.com/oono/ApplicableMath/ApplicableMath_files/AMII-33.pdf`.

**1.18** [Application of the Sackur-Tetrode equation[12]]
The following data are all under 1 atm.
The melting point of mercury is 234.2 K and the heat of fusion is 2.33 kJ/mol. The absolute entropy of solid mercury just at the melting temperature is 59.9 J/K·mol. The entropy increase of liquid between the melting point and the boiling point is 26.2 J/K·mol. The boiling point is 630K and the evaporation heat is 59.3 kJ/mol.
(1) Calculate the absolute entropy of mercury gas just above the boiling point.
(2) Assuming that mercury vapor is a monatomic ideal gas, obtain Planck's constant. The reader may use the value of $k_B$.

**Solution**

(1) We obtain the absolute entropy of mercury gas at 1 atm just above the boiling point simply by adding required entropy increases:

$$S = 59.9 + \frac{2330}{234.2} + 26.2 + \frac{59300}{630} = 190.1 \text{ J/mol·K}.$$

(2) The translational entropy of an ideal gas is

$$S = Nk_B \left\{ \log \frac{V}{N} + \frac{3}{2} \log T + \frac{5}{2} + \frac{3}{2} \log \frac{2\pi m k_B}{h^2} \right\}.$$

With the aid of the ideal gas law $PV = Nk_BT$ this can be rewritten as

$$\log P = \frac{5}{2} \log T + \frac{5}{2} + \log \left\{ k_B^{5/2} \left( \frac{2\pi m}{h^2} \right)^{3/2} \right\} - \frac{S}{Nk_B}.$$

Here $S$ is the entropy in the vapor phase for which the classical ideal gas law holds. Therefore, if this $S$ is estimated correctly, then the above equation should be very accurate.

$$\log \left\{ k_B^{5/2} \left( \frac{2\pi m}{h^2} \right)^{3/2} \right\} = \frac{S}{Nk_B} + \log p - \frac{5}{2} \log T - \frac{5}{2} = 22.84 + 11.52 - 18.61 = 15.74$$

If we use the known parameter values, we get 15.82, an excellent agreement. $m = 0.20059/N_A$ and this gives $h = 6.81 \times 10^{-34}$ Js ($6.623 \times 10^{-34}$ Js is the accepted value).

**1.19** [Negative temperature]
Let us consider the two state spin system containing 1 mole of spins as discussed

---

[12] The best elementary exposition is found in F. Mandl, *Statistical Physics* (Wiley, 1971) Section 7.6.

in the text. Assume that under the magnetic field $B$, the energy gap between the up and down spin states is 600 K per spin. Suppose the initial temperature of the magnet is $-500K$.
(1) What is the temperature of this system measured with an ideal gas thermometer containing $10^{-6}$ moles of monatomic gas particles?
(2) If, instead, the original negative temperature system is thermally equilibrated with a 1 mole of ideal monatomic gas that is initially 200K, what is the equilibrium temperature?

**Solution**

(1) The relation between the magnetization and the temperature can be solved as

$$m = N_A \tanh\frac{300}{T}. \tag{6.1.96}$$

We have already computed the problem

$$300 N_A k_B \tanh\frac{300}{500} = -300 k_B N_A \tanh\frac{300}{T} + \frac{3}{2} n k_B T \tag{6.1.97}$$

In this case we may expect that the temperature is extremely high, so

$$300 N_A \tanh\frac{300}{500} = 161 N_A \simeq \frac{3}{2} nT \Rightarrow T = 1.07 \times 10^8 \text{K}. \tag{6.1.98}$$

Since the temperature is outrageously high, we must pay attention to relativity. That is, the gas must be superrelativistic. Then, is the temperature higher or lower? (You can of course get the answer quantitatively easily.)

(2)

$$161 N_A k_B + \frac{3}{2} 200 N_A k_B = -300 N_A k_B \tanh\frac{300}{T} + \frac{3}{2} N_A k_B T \tag{6.1.99}$$

That is, we must solve

$$361 = -300 \tanh\frac{300}{T} + \frac{3}{2} T \tag{6.1.100}$$

That is, $T = 374$K.

## 6.2 Problems for Chapter 2

**2.1** [On derivation of canonical distribution]
A textbook of statistical mechanics has the following derivation of the canonical distribution:
"The distribution function must not depend on the properties of particular systems, but must be universal. That is, the probability $P(E_{\mathrm{I}})$ (resp., $P(E_{\mathrm{II}})$) for the system I (resp., II) to have energy $E_{\mathrm{I}}$ (resp., $E_{\mathrm{II}}$) and the probability $P(E)$ for the compound system of I and II to have energy $E = E_{\mathrm{I}} + E_{\mathrm{II}}$ must have the same functional form. This must be so, as long as we expect statistical mechanics holds universally. Therefore,

$$P(E_{\mathrm{I}})P(E_{\mathrm{II}}) = P(E_{\mathrm{I}} + E_{\mathrm{II}}). \tag{6.2.1}$$

For this equation to be valid, we can prove that $P(E)$ must have the following functional form:

$$P(E) = Ce^{-\beta E}, \tag{6.2.2}$$

where $C$ is a constant."

Is this argument correct?

**Solution**
Here, we must interpret $C$ as the normalization constant. Therefore, $C$ must be a functional of $E$ (or rather, the system Hamiltonian). This implies that $P(E)$ is not a simple function of $E$ but a functional of the system Hamiltonian, so, although (6.2.2) is a solution, we cannot conclude from (6.2.1), even if we assume smoothness (recall Cauchy), that this is the only solution; we may conclude that $\log P$ is a linear functional of $H$, but no further restriction is possible.

Everyone knows that (6.2.1) is incorrect, generally speaking, since $\sum_{E_{\mathrm{I}}} P(E_{\mathrm{I}})P(E_{\mathrm{II}}) = P(E)$. Precisely speaking, we must say that (6.2.1) holds for the most probable partition of energy. As can be seen from the correct answer $P = e^{-\beta(E-A)}$,

$$P_{\mathrm{I}}(E_{\mathrm{I}})P_{\mathrm{II}}(E_{\mathrm{II}}) = P_{\mathrm{I+II}}(E_{\mathrm{I}} + E_{\mathrm{II}}) \tag{6.2.3}$$

holds for the most probable partition of energy.

**2.2** [Elementary problem about spin system]
Due to the ligand field the degeneracy of the $d$-orbitals of the chromium ion $\mathrm{Cr}^{3+}$ is lifted, and the spin Hamiltonian has the following form

$$H = D(S_z^2 - S(S+1)/2), \tag{6.2.4}$$

where $D > 0$ is a constant with $S = 3/2$ (the cation is in the term ${}^4\mathrm{F}_{3/2}$).
(0) Why can you apply statistical mechanics to this 'single' ion?
(1) Compute the occupation probability of each energy level at temperature $T$ (you may use the standard notation $\beta = 1/k_BT$).
(2) Calculate the entropy.
(3) At high temperatures approximately we have $C = k_B(T_0/T)^2$ with $T_0 = 0.18$ K. Determine $D$ in K.

**Solution**
(0) Statistical mechanics exploits the fact that any macroscopic system may be considered as a set of numerous statistically independent collection of subsystems. That is, if there is statistical independence and additivity, we may apply the statistical mechanics framework to the collection.
(1) There are 4 states but there are only two energy levels with $E = 3D/8$ and $-13D/8$. Therefore, $S_z = \pm 3/2$ is with $p = 1/2(1 + e^{2\beta D})$ and $S_z = \pm 1/2$ is with $p = 1/2(1 + e^{-2\beta D})$.
(2) The easiest method is to use the Shannon formula:[13]

$$S = -2k_B \left[\frac{1}{2(1+x)} \log \frac{1}{2(1+x)} + \frac{x}{2(1+x)} \log \frac{x}{2(1+x)}\right] = k_B \left\{\log[2(1+x)] - \frac{x}{1+x} \log x\right\}, \tag{6.2.5}$$

where $x = e^{2\beta D}$.
(3) Setting $x$ as above, we have

$$C = T\frac{dS}{dT} = -(2D\beta)\frac{dS}{d2D\beta} = -2D\beta\frac{dx}{d2D\beta}\frac{dS}{dx} = -2D\beta x\frac{dS}{dx} = k_B(2D\beta)^2\frac{x}{(1+x)^2}. \tag{6.2.6}$$

Therefore,

$$C = k_B(D/k_B)^2/T^2. \tag{6.2.7}$$

This implies $D/k_B = T_0$. That is, $D$ is 0.18K.

**2.3**. [Vapor pressure of silicon]
The chemical potential $\mu_s$ of the atom in a solid is essentially given by the binding energy $\Delta$ of atom in the solid: $\mu_s = -\Delta$. Obtain the formula for the equilibrium vapor pressure of solid, and estimate the vapor pressure at room temperature of

---

[13] There are several ways to compute entropy. If you know probability explicitly, the Shannon formula may be useful. In this case, you must not forget that the sum is over the *elementary events*. The microcanonical way is probably the least useful in practice. When you compute $S$ from the canonical ensemble, use $S = (E - A)/T$ with $E$ being calculated by the Gibbs-Helmholtz relation $\partial(A/T)/\partial(1/T) = -\partial \log Z/\partial\beta = E$.

silicon for which $\Delta = 3$eV.

**Solution**

This is a typical Physics 100 level question. We may assume that the gas is ideal, so its chemical potential is given by $\mu = k_BT\log(n/n_Q)$. The chemical potential of the atom in the solid is $-\Delta$. Therefore the equilibrium condition (the identity of chemical potentials in two phases) gives

$$n = n_Q e^{-\beta\Delta} \Rightarrow p = k_BTn \tag{6.2.8}$$

We know $\hat{m} = 28$, so

$$n_Q = 28^{3/2} \times 10^{30} \simeq 1.5 \times 10^{32}. \tag{6.2.9}$$

Therefore,

$$\begin{aligned} P &= k_BTn_Qe^{-\beta\Delta} = 1.38 \times 10^{-23} \times 300 \times 1.5 \times 10^{32} e^{-3/0.026}, &(6.2.10)\\ &= 4.8 \times 10^{-39} \text{ (Pa)}. &(6.2.11)\end{aligned}$$

**2.4** [Specific heat]

Suppose that a (3D) classical mechanical system has the following Hamiltonian

$$H = \sum_{i=1}^{N} a_k|\boldsymbol{p}_k + \boldsymbol{c}_k|^s \tag{6.2.12}$$

where $a_k$ ($k = 1, \cdots, N$), $s$ are positive constants, and $\boldsymbol{c}_k$ are constant 3-vectors. Without any explicit calculation compute the specific heat.

**Solution**

The partition function $Z$ reads

$$Z = \frac{V^N}{N!h^{3N}} \prod_k \int d^3p\, e^{-\beta a_k|\boldsymbol{p}+\boldsymbol{c}_k|^s}. \tag{6.2.13}$$

Frist of all, $\boldsymbol{c}_k$ may be ignored by shifting the integration ranges. You could use the scaled variable $\lambda$ defined by $\beta a_k p^s = \lambda^s$ for each integral. However, the author recommends you to use Dimensional Analysis.

A recommended solution begins here. $\beta a_k p^s$ is dimensionless, so

$$[p] = [\beta a_k]^{-1/s}. \tag{6.2.14}$$

On the other hand,

$$[Zh^{3N}/V^N] = [p]^{3N} \propto [\beta^{-3N/s}] = (k_BT)^{3N/s}. \tag{6.2.15}$$

Therefore, even the internal energy (the *Gibbs-Helmholtz relation*):

$$E = \frac{\partial \beta A}{\partial \beta} = -\frac{\partial \log \beta^{-3N/s}}{\partial \beta} \tag{6.2.16}$$

is independent of $a_k$, so is $C$.

From the fact that $a_k$ is factored out, without any calculation you may conclude that $C$ is independent of $a_k$.

Eq.(6.2.16) is correct, so you can proceed with this result, but let us proceed in a more conventional way (although this is not a recommended approach). (6.2.13) reads

$$Z = \frac{V^N}{N!h^{3N}} \prod_k (\beta a_k)^{-3/s} \int 4\pi\lambda^2 e^{-\lambda^s} d\lambda \propto \prod_k (\beta a_k)^{-3/s}. \tag{6.2.17}$$

The Gibbs-Helmholtz relation immediately tells us that $E$ is not dependent on $a_k$. $E = 3Nk_BT/s$ and $C = 3Nk_B/s$.

The potential energy of a permanent electric dipole $\boldsymbol{p}$ is $U = -\boldsymbol{p} \cdot \boldsymbol{E}$ in the electric field $\boldsymbol{E}$. Obtain the electric susceptibility of the system.

**Solution**

We must obtain the expectation value of the polarization $\boldsymbol{P}$ per unit volume of the gas. In this case we ignore the interaction among gas particles, so the partition function becomes a product of one particle partition functions. Furthermore, since the translational motion of the particles has nothing to do with the polarization, we have only to compute the canonical partition function for a single dipole:

$$z(\boldsymbol{E}) = \int d\boldsymbol{e}\, e^{\beta p \boldsymbol{e}\cdot\boldsymbol{E}} = \int d\boldsymbol{e}\, e^{\beta pE\cos\theta}, \tag{6.2.18}$$

where $\boldsymbol{e}$ is the directional unit vector of the dipole moment with respect to the electric field direction, $E = |\boldsymbol{E}|$, and the angle between $\boldsymbol{E}$ and $\boldsymbol{p}$ (or $\boldsymbol{e}$) is $\theta$. The integration is on the unit sphere and can be computed as

$$z(\boldsymbol{E}) = 2\pi \int d\theta\, \sin\theta\, e^{\beta pE\cos\theta} = 2\pi \int_{-1}^{1} dx\, e^{\beta pEx} = \frac{4\pi}{\beta pE} \sinh \beta pE. \tag{6.2.19}$$

From the structure of $z$ we can immediately see

$$\langle \boldsymbol{p} \rangle = k_BT \frac{\partial}{\partial \boldsymbol{E}} \log z(\boldsymbol{E}) = pL(\beta pE)\frac{\boldsymbol{E}}{E} \tag{6.2.20}$$

where $L(x)$ is the Langevin function

$$L(x) = \coth x - \frac{1}{x}. \tag{6.2.21}$$

$\langle \boldsymbol{p} \rangle$ times the number of particle per volume $n$ (= the number density) is the polarization $\boldsymbol{P} = n\langle \boldsymbol{p} \rangle$.

The correspondence to thermodynamics is as follows. $-k_BTN \log z(\boldsymbol{E})$ is, as $\boldsymbol{E}$ is written explicitly, not the Helmholtz free energy $A(\boldsymbol{P})$, but its Legendre transformation $\Phi(\boldsymbol{E}) = A(\boldsymbol{P}) - \boldsymbol{P} \cdot \boldsymbol{E}$. That is,

$$dA = -SdT + \boldsymbol{E} \cdot d\boldsymbol{P}, \quad d\Phi = -SdT - \boldsymbol{P} \cdot d\boldsymbol{E}. \tag{6.2.22}$$

Therefore,

$$\boldsymbol{P} = -\frac{\partial}{\partial \boldsymbol{E}} \Phi(\boldsymbol{E}) = k_BTN \frac{\partial}{\partial \boldsymbol{E}} \log z(\boldsymbol{E}), \tag{6.2.23}$$

which is equivalent to (including the correct sign) (6.2.20).

To obtain the susceptibility (dielectric constant), we must differentiate $\boldsymbol{P}$ with respect to $\boldsymbol{E}$. It is a diagonal matrix proportional to the unit $3 \times 3$ matrix $I$ as

$$\varepsilon = \beta p^2 N L'(\beta p N) I. \tag{6.2.24}$$

**2.6** [Internal degree of freedom]

There is a classical ideal gas of volume $V$ consisting of $N$ molecules whose internal degrees of freedom are expressed by a single (quantum-mechanical) harmonic oscillator with a frequency $\nu$. The frequency depends on the volume of the system as

$$\frac{d\log \nu}{d\log V} = \gamma. \tag{6.2.25}$$

(1) Obtain the pressure of the system.

(2) Obtain the constant pressure specific heat $C_P$.

**Solution**

(1) The partition function is given by

$$Z = Z_0 Z_I, \tag{6.2.26}$$

where $Z_0$ is the canonical partition function for the classical ideal gas, and $Z_I$ is the partition function for the internal harmonic degree of freedom. You may use the already computed results:

$$Z_0 = \frac{V^N}{N!} \left( \frac{2mk_BT}{h^2} \right)^{3N/2} \tag{6.2.27}$$

and

$$Z_I = \left( \frac{1}{2\sinh \beta h\nu/2} \right)^N. \tag{6.2.28}$$

Therefore

$$P = -\left.\frac{\partial A}{\partial V}\right|_T = \frac{Nk_BT}{V} + P_I, \tag{6.2.29}$$

where

$$P_1 = \left.\frac{\partial k_BT \log Z_I}{\partial V}\right|_T = -Nk_BT\frac{\cosh \beta h\nu/2}{\sinh \beta h\nu/2}\frac{\beta h}{2}\frac{\partial \nu}{\partial V} = -\frac{Nh\nu}{2V}\gamma \coth\frac{\beta h\nu}{2}. \tag{6.2.30}$$

Notice that if we write the total energy of the internal degrees of freedom

$$U = \frac{Nh\nu}{2}\coth\frac{\beta h\nu}{2} = \frac{h\nu}{2} + \frac{h\nu}{e^{\beta h\nu}-1}, \tag{6.2.31}$$

then

$$P_I = -\gamma\frac{U}{V}. \tag{6.2.32}$$

That is,

$$P = \frac{Nk_BT}{V} - \gamma\frac{U}{V}. \tag{6.2.33}$$

(2) The best way is to use enthalpy $H = E + PV$:

$$C_P = \left.\frac{\partial H}{\partial T}\right|_P. \tag{6.2.34}$$

Again, you may use the results you know (you may assume there is only translational and oscillatory degrees of freedom): the internal energy is

$$E = \frac{3}{2}Nk_BT + \frac{Nh\nu}{2}\coth\frac{\beta h\nu}{2}. \tag{6.2.35}$$

Therefore,

$$H = \frac{5}{2}Nk_BT + (1-\gamma)U. \tag{6.2.36}$$

Use (6.2.31).

$$C_P = \frac{5}{2}Nk_B + (1-\gamma)Nk_B\left(\frac{\beta h\nu}{e^{\beta h\nu}-1}\right)^2 e^{\beta h\nu}. \tag{6.2.37}$$

**2.7** [Application of equipartition of energy][14]

---

[14] original credit: B. Widom

The internal motion of some ring puckering molecules (e.g., cyclobutanone) can be described by the following Hamiltonian:

$$H = \frac{p^2}{2m} + ax^4, \tag{6.2.38}$$

where $m$ is the effective mass of the oscillator and $a$ is a positive constant. Obtain the constant volume specific heat of this gas around the room temperature. Do not forget the contribution of translational degrees of freedom, etc.

**Solution**

We use the equipartition of energy:

$$2\left\langle \frac{p^2}{2m} \right\rangle = k_BT. \tag{6.2.39}$$

and

$$4\left\langle ax^4 \right\rangle = k_BT. \tag{6.2.40}$$

Therefore,

$$\langle H \rangle = \frac{3}{4}k_BT. \tag{6.2.41}$$

We must pay attention to the translational and rotational degrees of freedom. This gives $3k_BT$. Therefore,

$$C_V = \frac{15}{4}k_B \tag{6.2.42}$$

per molecule.

**2.8**. [Equipartition of energy for relativistic gas]

For a classical relativistic ideal gas show that the counterpart of the law of equipartition of kinetic energy reads

$$\left\langle \frac{1}{2}\frac{mv_x^2}{\sqrt{1-v^2/c^2}} \right\rangle = \frac{1}{2}k_BT. \tag{6.2.43}$$

**Solution**

The total energy of the particle $\varepsilon$ is obtained from the Lorentz invariance $\boldsymbol{p}^2 - \varepsilon^2/c^2 = -m^2c^2$ as

$$\varepsilon = c\sqrt{\boldsymbol{p}^2 + m^2c^2}. \tag{6.2.44}$$

We know the general relation to demonstrate the equipartition:

$$\left\langle p_x \frac{\partial \varepsilon}{\partial p_x} \right\rangle = \langle p_x v_x \rangle = k_BT. \tag{6.2.45}$$

Note that

$$p_i = \frac{mv_i}{\sqrt{1-(v/c)^2}}, \tag{6.2.46}$$

This concludes the demonstration.

**2.9** [An equality about canonical ensemble]
Let $\Phi$ be the total potential energy of classical system. Show

$$\langle \Delta\Phi \rangle = \beta \left\langle (\nabla\Phi)^2 \right\rangle . \tag{6.2.47}$$

Here, the Laplacian and the nabla are understood as operators in the 3N-space.

**Solution**
Let $H = K + \Phi$, where $K$ is the total kinetic energy.

$$\begin{aligned} \langle \Delta\Phi \rangle &= \frac{1}{Z}\int d\Gamma\,(div\,grad\Phi)e^{-\beta(K+\Phi)}, \\ &= \frac{1}{Z}\int d\Gamma\,div\left(e^{-\beta H}\nabla\Phi\right) - \frac{1}{Z}\int d\Gamma\,\nabla\Phi\cdot\nabla e^{-\beta H}, \\ &= -\frac{1}{Z}\int d\Gamma\,\nabla\Phi\cdot\nabla e^{-\beta H} = \beta\langle(\nabla\Phi)^2\rangle. \end{aligned}$$

The contribution from the boundary may be ignored thanks to the Boltzmann factor.

**2.10** [Density operator for free particles: perhaps an elementary QM review]
The canonical density operator is given by

$$\rho = \frac{1}{Z}e^{-\beta H}, \tag{6.2.48}$$

where $H$ is the system Hamiltonian and $Z$ is the canonical partition function. Let us consider a single particle confined in a 3D cube of edge length $L$. We wish to compute the position representation of the density operator $\langle x|\rho|x'\rangle$ ($x$ and $x'$ are 3D position vectors, and bras and kets are normalized).

Let $U(\beta) = e^{-\beta H}$ and $H = p^2/2m$. There are two ways to compute $\langle x|U(\beta)|x'\rangle$:
A.
(1) Show that

$$\frac{\partial}{\partial\beta}\langle x|U(\beta)|x'\rangle = \frac{\hbar^2}{2m}\Delta_x\langle x|U(\beta)|x'\rangle, \tag{6.2.49}$$

where $\Delta_x$ is the Laplacian with respect to the coordinates $x$.
(2) What is the initial condition (i.e., $\langle x|U(0)|x\rangle$)?

(3) Solve the equation in (1) with the correct initial condition. You may use a simple boundary condition assuming the volume is very large (and temperature is not too low).
(4) Compute $Z$, using the result in (3). You may use (3) to study the finite volume system as long as the temperature is not too low.

B.
We can directly compute $\langle x|U(\beta)|x'\rangle$ with the aid of the momentum representation of $U(\beta)$:

$$\langle p|U(\beta)|p'\rangle = e^{-\beta p^2/2m}\delta(p-p').$$

(5) We use

$$\langle x|U(\beta)|x'\rangle = \int d^3p\, d^3p' \langle x|p\rangle\langle p|U(\beta)|p'\rangle\langle p'|x'\rangle.$$

What is $\langle x|p\rangle$? You may assume the infinite volume normalization (i.e., the $\delta$-function normalization: $\langle p|p'\rangle = \delta(p-p')$).
(6) Perform the integral in (5).

**Solution**
A(1) We immediately obtain

$$-\frac{d}{d\beta}U = HU, \tag{6.2.50}$$

so its position representation is obtained as given. Notice that

$$\langle x|H|x'\rangle = -\frac{\hbar^2}{2m}\Delta_x\delta(x-x'). \tag{6.2.51}$$

(2) $U(0) = 1$, so $\langle x|U(0)|x\rangle = \delta(x-x')$ (if you use the continuum approximation) or $= \delta_{x,x'}$ (if you honestly treat the finiteness of the system).
(3) This is a diffusion equation, so the solution may be obtained by looking up any standard textbook; it is the Green's function with the vanishing boundary condition at infinity

$$\langle x|U(\beta)|x'\rangle = \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2} e^{-mk_BT(x-x')^2/2\hbar^2}.$$

This clearly exhibits that quantum effect becomes important at low temperatures (as can easily be guessed from the thermal wave length proportional to $1/\sqrt{T}$).
(4) $Z = TrU(\beta)$, so

$$Z = \int d^3x\langle x|U(\beta)|x\rangle = \int d^3x \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2} = V\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}.$$

This is consistent with (2.2.3).

(5) $|p\rangle$ is an eigenket of $H$ belonging to the eigenvalue $p^2/2m$: $H|p\rangle = (p^2/2m)|p\rangle$. Therefore,

$$-\frac{\hbar^2}{2m}\Delta_x\langle x|p\rangle = \frac{p^2}{2m}\langle x|p\rangle.$$

The boundary condition is a periodic boundary condition. The equation is essentially the harmonic oscillator equation, so the solution must be proportional to $e^{ipx/\hbar}$. We need a normalization condition:

$$\int d^3x e^{i(p-p')x/\hbar} = h^3\delta(p-p'),$$

so

$$\langle x|p\rangle = \frac{1}{h^{3/2}}e^{ipx/\hbar}.$$

(6)

$$\langle x|U(\beta)|x'\rangle = \int d^3p\,\langle x|p\rangle e^{-\beta p^2/2m}\langle p|x'\rangle = \int d^3p\,e^{-\beta p^2/2m+i(x-x')p/\hbar}.$$

This is a simple Gaussian integral, so indeed the answer agrees with (3) above.

**2.11** [Density operator for a spin system]

Let $\rho$ be the density operator of a single 1/2 quantum spin system whose Hamiltonian[15] is given by $H = -\gamma\boldsymbol{\sigma}\cdot\boldsymbol{B}$, where $\boldsymbol{\sigma}$ is $(\sigma_x, \sigma_y, \sigma_z)$ in terms of the Pauli spin operators.

(1) Obtain the matrix representation of $\rho$ that diagonalizes $\sigma_z$.

(2) Find the average of $\sigma_y$.

(3) Obtain the matrix representation of $\rho$ that diagonalizes $\sigma_x$.

**Solution**

(1) We take the direction of $\boldsymbol{B}$ to be the $z$ axis.

$$\rho = \frac{1}{C}\begin{pmatrix} e^{\beta\gamma B} & 0 \\ 0 & e^{-\beta\gamma B}\end{pmatrix}, \tag{6.2.52}$$

where $C$ is the normalization constant: the trace of the matrix in the above formula, so $C = 2\cosh\beta\gamma H$.

If you wish to do the original problem we need the following calculation. Notice that $(\boldsymbol{n}\cdot\boldsymbol{\sigma})^2 = I$, where $\boldsymbol{n}$ is a unit vector.

$$e^{\gamma\boldsymbol{B}\cdot\boldsymbol{\sigma}} = \sum_{n=0}^{\infty}\frac{1}{(2n)!}(\gamma B)^{2n} + \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(\gamma B)^{2n+1}\frac{\boldsymbol{B}}{B}\cdot\boldsymbol{\sigma} = \cosh\gamma B + \frac{\boldsymbol{B}}{B}\cdot\boldsymbol{\sigma}\sinh\gamma B. \tag{6.2.53}$$

---

[15] Precisely speaking, this is te interaction Hamiltonian of the system with the magnetic field.

(2)

$$\langle \sigma_y \rangle = Tr\sigma_y\rho = Tr \begin{pmatrix} 0 & ie^{\beta\gamma B} \\ -ie^{-\beta\gamma B} & 0 \end{pmatrix} = 0. \tag{6.2.54}$$

This should be obvious without any calculation.

(3) With the basis that diagonalizes $\sigma_z$ we have

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{6.2.55}$$

Therefore

$$\sigma_x \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}. \tag{6.2.56}$$

That is, the following orthogonal (actually, unitary as well) matrix:

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \tag{6.2.57}$$

diagonalizes $\sigma_x$ as $U^*\sigma_x U$. Therefore,

$$U^* \begin{pmatrix} 0 & ie^{\beta\gamma B} \\ -ie^{-\beta\gamma B} & 0 \end{pmatrix} U = \frac{1}{2} \begin{pmatrix} 1 & -\tanh\beta\gamma B \\ -\tanh\beta\gamma B & 1 \end{pmatrix}. \tag{6.2.58}$$

**2.12** [Legendre vs Laplace]
Consider an ideal gas consisting of $N$ atoms under constant pressure $P$ and temperature $T$.
(1) What is the most convenient partition function and the thermodynamic potential? Compute the partition function. You may use the ideal gas canonical partition function.
(2) Obtain the enthalpy of the system.

**Solution**
(1) We should use the pressure ensemble:

$$Q = \int dV Z(T, V) e^{-\beta PV}. \tag{6.2.59}$$

This is related naturally to the Gibbs free energy: $A + PV = G = -k_BT\log Q$. (Since $N \gg 1$, you need not worry about $\pm 1$ in $N$.)

$$Q = \int_0^\infty dV \frac{1}{N!}\left(\frac{2\pi m k_BT}{h^2}\right)^{3N/2} V^N e^{-\beta PV} = \left(\frac{2\pi m k_BT}{h^2}\right)^{3N/2}(\beta P)^{-N}. \quad (6.2.60)$$

This gives

$$G = -Nk_BT\log\left(\frac{2\pi m k_BT}{h^2}\right)^{3/2} + Nk_BT\log\frac{P}{k_BT}, \quad (6.2.61)$$

which is

$$G = A - k_BT\log N! + Nk_BT\log\frac{PV}{k_BT} = A + Nk_BT = A + PV. \quad (6.2.62)$$

Consistent!

(2) We can use an analogue of the Gibbs-Helmholtz relation

$$H = \frac{\partial G/T}{\partial 1/T} = -\frac{\partial \log Q}{\partial \beta} = \frac{5}{2}Nk_BT. \quad (6.2.63)$$

Thermodynamically, $H = E + PV = E + Nk_BT$. Consistent. This is of course consistent with the constant pressure specific heat.

**2.13** [Constant magnetic field ensemble]

The following situation is the same as **1.18**: $N$ lattice sites have spins of $S = 1$ (in the term ${}^3P$), and the spin Hamiltonian has the following form:

$$H = DS_z^2. \quad (6.2.64)$$

(1) Consider this as the constant magnetic field ($B$) ensemble (also constant $T$ is assumed), and construct the corresponding generalized canonical partition function. The magnetization is given by $M = \mu\sum S_{zi}$.

(2) Compute the magnetization as a function of the magnetic field and temperature.

**Solution**

(1) Since

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{B}{T}dM + \cdots, \quad (6.2.65)$$

the generalized partition function reads

$$Z_H = (1 + e^{-\beta D + \beta\mu B} + e^{-\beta D - \beta\mu B})^N. \quad (6.2.66)$$

(2) The resulting generalized Massieu function

$$\psi = k_B \log Z_H \tag{6.2.67}$$

satisfies

$$d\psi = -Ed\left(\frac{1}{T}\right) + Md\left(\frac{B}{T}\right) + \cdots. \tag{6.2.68}$$

Therefore,

$$M = \frac{\partial \psi}{\partial (B/T)} = N\frac{\mu(e^{-\beta D+\beta\mu B} - e^{-\beta D-\beta\mu B})}{1 + e^{-\beta D+\beta\mu B} + e^{-\beta D-\beta\mu B}} = N\frac{\sinh \beta\mu B}{e^{-\beta D}/2 + \cosh \beta\mu B}. \tag{6.2.69}$$

**2.14** [Absorption of mixed ideal gas, or convenient partition function]
There is a gas mixture consisting of two distinct atomic species $A$ and $B$. The mixture is an ideal gas and the partial pressures of $X$ is $p_X(X = A$ or $B)$. The gas is in equilibrium with an adsorbing metal surface on which there are adsorption sites. Atom X adsorbed at the site is with energy $-E_X$ on the average relative to the one in the gas phase, where $X = A$ or $B$. Each surface site can accommodate at most one atom. Assume that [Hint: I assume that you know how to calculate the chemical potentials of the atoms, knowing the molecular weights.]
(1) Write down the 'partition function' (use the most convenient one) for the single site.
(2) Obtain the average surface concentration $n_X$ ($X = A$ or $B$) of atoms $A$ and $B$.
(3) Under a given (partial) pressures of $A$ and $B$ $n_E : n_A : n_B = 1 : 1 : 18$ (here $E$ means empty). Find the maximum concentration $n_A$ obtainable with changing only the partial pressure of B. (UIUC Qual F95).

**Solution**
(1) Each adsorption site has three states, empty, occupied by $A$ and occupied by $B$. Therefore, for a single site the grand partition function function reads

$$\Xi = 1 + e^{\beta(E_A+\mu_A)} + e^{\beta(E_B+\mu_B)}, \tag{6.2.70}$$

where $\mu_X$ is the chemical potential of $X$ that can be written as

$$\mu_X = k_BT \log(\beta p_X/n_{QX}) = k_BT \log(p_X/p_{QX}) \tag{6.2.71}$$

Here, $n_{QX}$ is the 'quantum density' depending on $T$ and the mass (see **1.4.5**), and $p_{QX} = n_{QX}k_BT$ may be called the 'quantum pressure.' Therefore,

$$\Xi = 1 + p_Ae^{\beta E_A}/p_{QA} + p_Be^{\beta E_B}/p_{QB}. \tag{6.2.72}$$

You need not compute $n_Q$ that is a matter of simple calculation as in **1.4.5**.
(2) As seen in **6-2** we may write

$$d(k_BT\log\Xi) = SdT + pdV + N_A d\mu_A + N_B d\mu_B + \cdots, \tag{6.2.73}$$

so if we apply this to each site $N_X$ may be interpreted as $n_X$:

$$\begin{aligned} n_A &= \frac{\partial(k_BT\log\Xi)}{\partial\mu_A} = \frac{p_A e^{\beta E_A}/p_{QA}}{1+p_A e^{\beta E_A}/p_{QA}+p_B e^{\beta E_B}/p_{QB}}, && (6.2.74)\\ n_B &= \frac{p_B e^{\beta E_B}/p_{QB}}{1+p_A e^{\beta E_A}/p_{QA}+p_B e^{\beta E_B}/p_{QB}}, && (6.2.75)\\ n_E &= \frac{1}{1+p_A e^{\beta E_A}/p_{QA}+p_B e^{\beta E_B}/p_{QB}},. && (6.2.76) \end{aligned}$$

(3) It should be intuitively obvious that $A$ absorption is facilitated if there is no competition. Hence, the maximum concentration should be accomplished by $p_B = 0$. Now, $n_A = n_E$ for a given $p_B$, so we must conclude that $p_A e^{\beta E_A}/p_{QA} = 1$. Therefore, the max conc must be 0.5.

**2.15** [Absorption on catalytic surface]
There are $N$ absorption centers on the catalyst surface exposed to a gas (ideal gas) of a certain chemical. Each absorption center can accommodate at most two particles. The partition function for the single particle absorption state is $a_1$ and the two particle absorption state is $a_2$.
(1) Write down the single site (grand) partition function.
(2) Let $a_1 = 0$ (i.e., absorption is possible only when a pair is formed). The average number of particles absorbed on the catalytic surface is $n_0$. Find the chemical potential of the particles.
(3) Now, the pressure of the chemical is doubled (with the temperature kept constant) and the average number of particles absorbed on the catalytic surface is $n_1$. Find $n_1$ in terms of $N$ and $n_0$. $a_1$ is still assumed to be zero.
(4) If $a_1 > 0$, does the number of absorbed molecules increase from $n_0$ in (2) (i.e., the $a_1 = 0$ case)? Demonstrate your answer and give a brief physical explanation.

**Solution**
(1)

$$\Xi_{single} = 1 + a_1 e^{\beta\mu} + a_2 e^{2\beta\mu}. \tag{6.2.77}$$

(2) The average particle numbers in the system (= the catalytic surface; notice that the gas phase is treated as a chemical reservoir) is obtained by $\partial\log\Xi/\partial(\beta\mu)$

$(\Xi = \Xi_{single}^N)$:

$$n_0 = N\frac{2a_2x^2}{1+a_2x^2}, \tag{6.2.78}$$

where $x = e^{\beta\mu}$ (called *fugacity*). Therefore, we have two possibilities, but $x$ cannot be negative, since $\beta\mu \in \boldsymbol{R}$:

$$x = \sqrt{\frac{n_0/N}{(2-n_0/N)a_2}}. \tag{6.2.79}$$

Therefore,

$$\mu = k_BT \log \sqrt{\frac{n_0/N}{(2-n_0/N)a_2}}. \tag{6.2.80}$$

(3) If the pressure is doubled, the fugacity doubles. Therefore, (6.2.79) tells us that

$$2\sqrt{\frac{n_0/N}{(2-n_0/N)a_2}} = \sqrt{\frac{n_1/N}{(2-n_1/N)a_2}}. \tag{6.2.81}$$

Solving this for $n_1$, we get

$$n_1 = \frac{8n_0}{2+3n_0/N}. \tag{6.2.82}$$

(4) Equation (6.2.78) now reads

$$n = N\frac{a_1x+2a_2x^2}{1+a_1x+a_2x^2}. \tag{6.2.83}$$

This may be written as

$$n = N\left(1+\frac{a_2x^2-1}{1+a_1x+a_2x^2}\right). \tag{6.2.84}$$

$a_1x > 0$ implies that if $a_2x^2 > 1$ (i.e., $n/N > 1$), then increasing $a_1$ (that is, favoring monomers) decreases the number of the absorbed molecules; otherwise, opposite.

Physically (or intuitively), this should be natural, because if monomers are favored when dimers are also sufficiently favored, they compete the sites. Thus, the double occupancy fraction decreases, so does the total number of absorbed molecules.

**2.16** [Gas under a weight]
Suppose there is a vertical cylindrical container of cross section $s$ whose top wall is a movable piston of cross section $s$ with mass $M$. The piston is assumed to move only in the vertical direction ($z$-direction) and feels gravity. The container contains

$N$ ($\gg 1$) classical noninteracting particles with mass $m$.
(1) Write down the Hamiltonian of the gas + piston system (write the piston vertical momentum as $p_M$).
(2) Obtain the pressure $P$ of the gas, and write the Hamiltonian in terms of $P$ and the volume of the gas $V = sz$, where $z$ is the position of the piston from the bottom of the container.
(3) Now, the mechanical variables are the phase variables of the gas system and the piston momentum $p_M$ and $z = V/s$. Compute the canonical partition function of the whole system.
(4) You should have realized that the calculation in (3), apart from the unimportant contribution in the thermodynamic limit of the piston momentum, is the calculation of the pressure ensemble. [That is, the heavy piston acts as a constant pressure device.] Obtain the equation of state of the gas in the cylinder (a trivial question).

**Solution**

(1)

$$H = \sum_{i=1}^{N} \frac{{p_i}^2}{2m} + \frac{p_M^2}{2M} + Mgz. \tag{6.2.85}$$

(2) From the force balance, we have

$$Ps = Mg \ \Rightarrow \ PV = Mgz. \tag{6.2.86}$$

Therefore, (6.2.85) can be rewritten as

$$H = \sum_{i=1}^{N} \frac{{p_i}^2}{2m} + \frac{p_M^2}{2M} + PV. \tag{6.2.87}$$

(3)

$$\begin{aligned} Z &= \frac{1}{N!h^{3N+1}} \int d^N\boldsymbol{p} d^N\boldsymbol{q} dp_M dz \, e^{-\beta H}, & (6.2.88)\\ &= \frac{1}{N!} \left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2} \left(\frac{2\pi M k_B T}{h^2}\right)^{1/2} \int V^N e^{-\beta PV} \frac{dV}{s} & (6.2.89)\\ &= \left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2} \left(\frac{2\pi M k_B T}{h^2}\right)^{1/2} (\beta P)^{-N}. & (6.2.90) \end{aligned}$$

Notice that this is almost the pressure ensemble for the ideal gas.

(4) We know (see **2.12**) the above $Z$ is proportional to the pressure ensemble partition function $Q$:

$$Z = Q\left(\frac{2\pi M k_B T}{h^2}\right)^{1/2}, \tag{6.2.91}$$

where

$$Q = \left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2}(\beta P)^{-N}. \tag{6.2.92}$$

We know from the Laplace-Legendre correspondence that $G = -k_B T \log Q$:

$$G = -Nk_BT\log\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} + Nk_BT\log\frac{P}{k_BT}, \tag{6.2.93}$$

$dG = -SdT + VdP$, so

$$V = \left.\frac{\partial G}{\partial P}\right|_T = \frac{Nk_BT}{P}. \tag{6.2.94}$$

This is the equation of state as expected. The enthalpy of the gas can be obtained by the Gibbs-Helmholtz relation

$$H = \left.\frac{\partial(G/T)}{\partial(1/T)}\right|_P = \frac{5}{2}Nk_BT. \tag{6.2.95}$$

**2.17** [Ideal gas with the aid of grand canonical ensemble]
Let us study the classical ideal gas with the aid of the grand canonical ensemble. Let $\mu$ be the chemical potential.
(1) Compute the grand canonical partition function for a monatomic ideal gas. Assume that the mass of the atom is $m$.
(2) Find the internal energy and the pressure as a function of chemical potential $\mu$.
(3) Suppose the expectation value of the number of particles is $N$. How is the chimical potentialdetermined?
(4) Are the results obtained above (especially the results of (2)) consistent with what you already know?

**Solution**
(1) By definition

$$\Xi = \sum_{N=0}^{\infty}\frac{1}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2}V^N e^{b\mu N} = \exp\left[\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2}Ve^{b\mu}\right] \tag{6.2.96}$$

(2) From this we get

$$P = \frac{k_B T}{V}\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} V e^{\beta\mu}. \tag{6.2.97}$$

Since

$$d(PV/T) = -Ed(1/T) + (P/T)dV + Nd(\mu/T), \tag{6.2.98}$$

$$E = T^2 \left.\frac{\partial k_B \log \Xi}{\partial T}\right|_{\mu/T,V} = \frac{3}{2}k_B T\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} V e^{\beta\mu}. \tag{6.2.99}$$

Noice that this is equal to $3PV/2$.
(3) Since

$$N = \frac{1}{k_B}\frac{\partial}{\partial\beta\mu}\frac{PV}{T} = \frac{\partial}{\partial\beta\mu}\log\Xi = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} V e^{\beta\mu}, \tag{6.2.100}$$

we obtain

$$\mu = k_B T \log\frac{N}{V}\left(\frac{2\pi m k_B T}{h^2}\right)^{-3/2}. \tag{6.2.101}$$

The result agrees with the result obtained in the text. If $n \ll n_Q$ (i.e., the classical case without overlapping of de Broglie wave packets), $\mu$ deviates verymuch to the negative side ($\mu \ll 0$).
(4) Thus, as has already been stated at various places, the results of the grand canonical ensemble completely reproduces the properties of the classic ideal gas.

**2.18** [To obtain the microcanonical partition function with the aid of Laplace inverse transformation]
Starting from

$$Z = \frac{V^N}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2}, \tag{6.2.102}$$

obtain the microcanonical partition fucntion $w(E,V)$ (with the aid of Laplace inverse transformation).
**Solution**

$$w(E,V) = \frac{1}{2\pi i}\int_{\beta^*-9\infty}^{\beta^*+i\infty} Z(\beta)e^{\beta E}d\beta. \tag{6.2.103}$$

If you can demonstrate the following formula, you may use it:

$$\frac{1}{2\pi i}\int_{s'-i\infty}^{s'+i\infty}\frac{e^{sx}}{x^{k+1}}ds = \frac{x^k}{\Gamma(k+1)}. \tag{6.2.104}$$

If $k$ is an integer, this is almost self-evident with the aid of Goursat's theorem, but in our case $k$ need not be an integer.

**2.19** [Equivalence of canonical and grand canonical ensembles]
Let us check the equivalence of grand canonical and canonical ensembles. That is, if we compute thermodynamic quantities in the thermodynamic limit, both give the same answers. Even experimentalists should look at this proof at least once in their lives.

The grand partition function $\Xi(T,\mu)$ and canonical partition function $Z(T,N)$ (the ground state energy is taken to be the origin of energy) are related as

$$\Xi(T,\mu) = \sum_{N=0}^{\infty} Z(T.N)e^{\beta\mu N}.$$

Let us assume that the system consists of $N$ (which is variable) particles in a box of volume $V$ and the total interaction potential $\Phi$ among particles is bounded from below by a number proportional to the number of particles $N$ in the system: $\Phi \geq -NB$, where $B$ is a (positive) constant. (The system Hamiltonian generally has the form of $H = K + \Phi$, where $K$ is the kinetic energy.)

Through answering the following almost trivial questions, we can demonstrate the ensemble equivalence (rigorously).
(1) Show that there is a constant $a$ such that

$$Z(T,N) \leq \left(\frac{aV}{N}\right)^N. \tag{6.2.105}$$

Actually, show (classically)

$$Z(T,N) \leq Z_0(T,N)e^{\beta NB},$$

where $Z_0$ is the canonical partition function for the ideal gas (e.g., (1.7.3)). This is just eq.(6.2.105) above
(2) Show that the infinite sum defining the grand partition function actually converges. The reader may use eq.(6.2.105) and $N! \sim (N/e)^N$ freely.
(3) Choose $N_0$ so that

$$\sum_{N=N_0}^{\infty} Z(T,N)e^{\beta\mu N} < 1.$$

Show that this $N_0$ may be chosen to be proportional to $V$ (that is, $N_0$ is at most extensive).
(4) Show the following almost trivial bounds:

$$\max_N Z(T,N)e^{\beta\mu N} \leq \Xi(T,\mu) \leq (N_0+1)\max_N Z(T,N)e^{\beta\mu N}.$$

(5) We are almost done, but to be explicit, show that $PV/Nk_BT$ obtained thermodynamically from the canonical partition function and that directly obtained from the grand partition function agree.

**Solution**
(1) The canonical partition function reads

$$Z(T,N) = \frac{1}{N!}\int d\Gamma e^{-\beta(K+\Phi)} \le \frac{1}{N!}\int d\Gamma e^{-\beta K} e^{\beta BN} = Z_0(T,N)e^{\beta NB}, \quad (6.2.106)$$

where $Z_0$ is the canonical partition function of the ideal gas. We know the kinetic part may be factorized into the individual particle contributions, and $N! \sim (N/e)^N$, so there must be $a$ satisfying the inequality.

**Remark**. The estimate is also correct quantum mechanically, so our proof being checked here is quite general.

(2) The grand partition function is a positive term series, and each term is bounded by the estimate in (1), so

$$\Xi(T,\mu) = \sum_{N=0}^{\infty} Z(T.N)e^{\beta\mu N} \le \sum_{N=0}^{\infty}\left(\frac{aV}{N}\right)^N e^{\beta\mu N} = \sum_{N=0}^{\infty}\left(\frac{aVe^{\beta\mu}}{N}\right)^N.$$

That is, with the aid of Stirling's formula,

$$\Xi(T,\mu) \le \sum_{N=0}^{\infty}\frac{1}{N!}(aVe^{\beta\mu-1})^N = \exp\left(aVe^{\beta\mu-1}\right).$$

The grand canonical partition function is a sum of positive terms, and bounded from above, so it must converge to a positive number.

For many realistic systems the interaction potentials have sufficiently hard repulsive cores, so the convergence is much quicker.

(3) This is the tail estimation to majorize it. Any crude choice will do, so we first 'overestimate' the sum beyond $N_0$ as

$$\sum_{N=N_0}^{\infty} Z(T,N)e^{\beta\mu N} \le \sum_{N=N_0}^{\infty}\frac{1}{N!}(aVe^{\beta\mu-1})^N \simeq \sum_{N=N_0}^{\infty}\left(\frac{aVe^{\beta\mu}}{N}\right)^N \quad (6.2.107)$$

Here, Stirling's formula has been used. For example, if we assume

$$\frac{aVe^{\beta\mu}}{N_0} < 0.1, \quad (6.2.108)$$

then

$$\sum_{N=N_0}^{\infty} Z(T,N)e^{\beta\mu N} < \sum_{N=N_0}^{\infty} 0.1^N. \tag{6.2.109}$$

The sum on the RHS is obviously bounded by 0.2 (by 1/9, at worst $N_0 = 1$). Thus, the choice (6.2.108) is enough. Such $N_0$ can clearly be chosen proportional to $V$.
(4) The grand partition function is a sum of positive terms, so it must be larger than any one term, especially larger than the largest term, in it:

$$\max_N Z(T,N)e^{\beta\mu N} \le \Xi(T,\mu). \tag{6.2.110}$$

Notice that the largest term cannot be less than 1, because the $N = 0$ term is never smaller than 1.[16] To obtain the upper bound $\Xi$ is divided into the sum up to $N_0 - 1$ and that beyond $N_0 - 1$:

$$\Xi(T,\mu) = \sum_{N=0}^{N_0-1} Z(T,N)e^{\beta\mu N} + \sum_{N=N_0}^{\infty} Z(T,N)e^{\beta\mu N}. \tag{6.2.111}$$

The second term on the right hand side is bounded by 1, which is not larger than the maximum term in the sum, so it is bounded by $\max_N Z(T,N)e^{\beta\mu N}$. Therefore,

$$\Xi(T,\mu) \le \sum_{N=0}^{N_0-1} Z(T,N)e^{\beta\mu N} + \max_N Z(T,N)e^{\beta\mu N}. \tag{6.2.112}$$

The sum in the above inequality must be less than the number of terms × the largest term:

$$\sum_{N=0}^{N_0-1} Z(T,N)e^{\beta\mu N} \le N_0 \max_N Z(T,N)e^{\beta\mu N}. \tag{6.2.113}$$

Therefore, we have

$$\Xi(T,\mu) \le (N_0+1) \max_N Z(T,N)e^{\beta\mu N}. \tag{6.2.114}$$

Combining this with (6.2.110) we get the desired result.

---

[16] Notice that $Z(T,0) \ge 1$: recall

$$Z(T,N) = \sum w(E,N)E^{-\beta E},$$

and $N = 0$ term is of course included, which is 1. This means $\max Z(T,N)e^{\beta\mu N} \ge 1$.

(5) The grand canonical ensemble asserts

$$\frac{PV}{Nk_BT} = \frac{1}{N}\log \Xi(T,\mu).$$

From the above inequality

$$\frac{1}{N}\log\left(\max_N Z(T,N)e^{\beta\mu N}\right) \le \frac{PV}{Nk_BT} \le \frac{1}{N}\log\left(\max_N Z(T,N)e^{\beta\mu N}\right) + \frac{1}{N}\log(N_0+1).$$

Notice that

$$\frac{1}{N}\log\max_N Z(T,N)e^{\beta\mu N} = \frac{1}{N}\max_N\{-\beta A + \beta\mu N\}$$

is a Legendre transformation of $A$ (recall $dA = -SdT - PdV + \mu dN$ or $A = -PV + \mu N$). Therefore, $\min_N\{A - \mu N\} = -PV$. This is the $PV$ obtained thermodynamically with the aid of the canonical ensemble results. That is,

$$\left(\frac{PV}{Nk_BT}\right)_{th} \le \frac{PV}{Nk_BT} \le \left(\frac{PV}{Nk_BT}\right)_{th} + \frac{1}{N}\log(N_0+1).$$

$\log(N_0+1)$ is bounded by a number proportional to $\log V$ as demonstrated in (3), so in the $N\to\infty$ limit the rightmost term behaves at worst as $(\log N)/N$, and may be ignored. Therefore, the pressure obtained thermodynamically from the Helmholtz free energy (obtained by the canonical ensemble) and the pressure directly obtained statistical mechanically with the aid of the grand canonical ensemble agree:

$$\left(\frac{PV}{Nk_BT}\right)_{th} = \frac{PV}{Nk_BT}.$$

**2.20** [Legendre transformation in convex analysis]

(1) We know that $-S$ is a convex function of internal energy $E$. Using the general property of the Legendre transformation, show that Helmholtz free energy $A$ is convex upward as a function of $T$. You may assume any derivative you wish to compute exists.

(2) When a phase transition occurs, the curve of $S(E)$ has a linear part as a function of $E$ (that is, $E$ can change under constant $T = T_e$). Then, $A$ as a function has a cusp at $T = T_e$ (that is, all the states corresponding to the flat part is collapsed to a point, the one-to-one nature of the Legendre transformation can be lost, if there is a phase transition). To illustrate this point, let us consider the following toy function

$$f(x) = \begin{cases} 2\tanh(x+1) - 2 & \text{for } x < -1, \\ 2x & \text{for } -1 \le x \le 1, \\ (x-1)^2 + 2x & \text{for } x > 1. \end{cases}$$

Sketch its Legendre transform $f^*(\alpha) = \sup_x[\alpha x - f(x)]$. [Do not try to compute the explicit formula.]

**Solution**

(1) $-S$ is a convex function of $E$, so that it is also a convex function of $-E$. We know

$$-\frac{A}{T} = \sup_T[-E/T - (-S)]. \tag{6.2.115}$$

Therefore, $-A/T$ is a convex function of $1/T$. A dirty way to proceed is to perform differentiation twice.

$$\frac{d(-A/T)}{d1/T} = -A - \frac{1}{T}\frac{dA}{d1/T} = -A + T\frac{dA}{dT}. \tag{6.2.116}$$

Therefore,

$$\frac{d^2(-A/T)}{d(1/T)^2} = -T^2\frac{d}{dT}\left(-A + T\frac{dA}{dT}\right) = -T^3\frac{d^2A}{dT^2} > 0 \tag{6.2.117}$$

Therefore, $A$ is a concave function of $T$.

(2) We consider the Legendre transformation

$$f^*(\alpha) = \sup_x[\alpha x - f(x)] \tag{6.2.118}$$

For $\alpha > 2$ this is easy, and we obtain $f^*(\alpha) = \alpha^2/2 - 1$. Between $-2$ and 2 of $x$ the slope of $f$ does not change and is 2, so $\alpha = 2$, which means $f^*(2) = 0$ is a cusp. For $\alpha < 2$, analytic calculation is not wise. We know $\alpha$ cannot be negative, and in the $\alpha \to 0$ limit, $f^* \to 4$, because $\lim_{x\to-\infty} f(x) = -4$. Since $f^*$ is convex, we can easily sketch its overall shape as below:

Remember that convex functions are continuous.

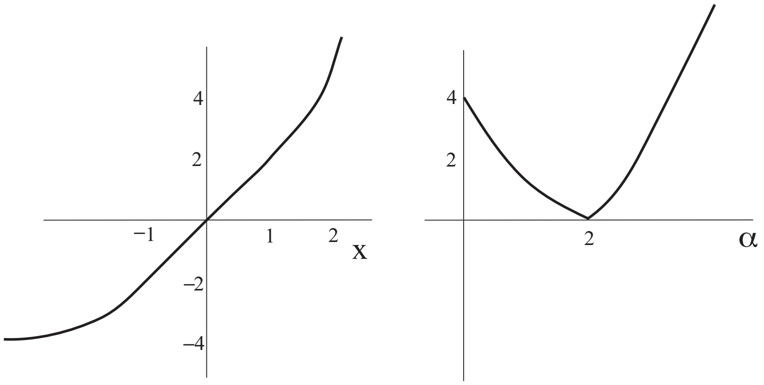
**2.21** [Information]

Suppose there are two fair dice. We assume that one dice is red and the other is green (that is, distinguishable). Let us record the numbers that are up in this order as $(n, m)$ $(n, m \in \{1, 2, \cdots, 6\})$.

(1) To know a particular pair of numbers $(a, b)$ unambiguously what information (in bits) do you need?

(2) You are told that the sum $a + b$ is not less than 5. What is the information you gain from this message?

(3) Next, you are told, one of the dice shows the face less than 3. What is the



**Left**: $f$, **Right**: $f^*$.

information you gain? (You must know the info obtained from (2) already.)
(4) Now, you are told that actually, the one of the dice in (3) is the red one. What is the information carried by this message?
(5) Finally, you are told that face pair is actually $(2, 5)$. What is the information in this final statement?

**Solution**

(1) There are 36 distinguishable states and they are all equally probable. Therefore, the total uncertainty is $\log_2 36 = 5.16$ bits, or the surprisal you have when you are told, say, $(1, 1)$ actually happens is 5.16 bits. That is, you need 5.16 bits of information to pinpoint a particular elementary event.
(2) There is no simpler way than actually to list all elementary states up: (1,1), (1,2), (1,3), (2,1), (2,2), (3, 1). These 6 states are excluded. Remaining are 30 states, all equally probably, so $\log_s 30 = 4.91$ bits is the uncertainty. That is, $5.16 - 4.91 = 0.25$ bits is the information in the message.
(3) Red = 1: Green = 4, 5 or 6
Red = 2: Green = 3, 4, 5 or 6.
Therefore, there are $7 \times 2 = 14$ states remaining. This uncertainty is $\log_2 14 = 3.81$. We had 4.91 bits of uncertainty, so this message must have conveyed 1.1 bits.
(4) Obviously, 1 bit.
(5) There is no uncertainty remaining, so 2.81 bits (this is, needless to say, the surprisal of an event of probability 1/7)

**2.22** [Variational principle for free energy (classical case)[17]]
Let $H = H_0 + V$ be a system Hamiltonian.
(1) Show that**q**

$$A \leq A_0 + \langle V \rangle_0,$$

---

[17] This holds quantum mechanically as well, but the proof is not this simple.

where $A$ is the free energy of the system with $H$ and $A_0$ that with $H_0$. $\langle\ \rangle_0$ is the average over the canonical distribution of the system with the Hamiltonian $H_0$. The inequality is (sometimes) called the Gibbs-Bogoliubov inequality.
(2) We can use the inequality to estimate $A$. If we can compute $A_0$ and $\langle V\rangle_0$ (that is the free energy for the system with $H_0$ and the average with respect to this system), then we can estimate the upper bound of $A$. Its minimum may be a good approximation to $A$. This is the idea of the variational approximation. Let us study an unharmonic oscillator with the Hamiltonian

$$H = \frac{1}{2m}p^2 + \frac{1}{2}kx^2 + \frac{1}{4}\alpha x^4,$$

where $m$, $k$ and $\alpha$ are positive constants. Let us define

$$H_0 = \frac{1}{2m}p^2 + \frac{1}{2}Kx^2.$$

Choose $K$ to obtain the best estimate of $A$ (you need not compute the estimate of $A$; it is easy but messy). You may use all the available results in the text.

**Solution**
(1)

$$\langle e^{-\beta V}\rangle_0 = \frac{1}{Z_0}\int d\Gamma e^{-\beta V} e^{-\beta H_0} = \frac{Z}{Z_0} = e^{-\beta(A-A_0)}.$$

Therefore, with the aid of Jensen's inequality

$$e^{-\beta\langle V\rangle_0} \le e^{-\beta(A-A_0)}.$$

That is, we are done.
(2) We know

$$A_0 = k_BT\log\left[\frac{\hbar\sqrt{K/m}}{k_BT}\right]$$

and (with the aid of $\langle x^4\rangle_0 = 3\langle x^2\rangle_0^2$ and equipartition of energy)

$$\left\langle \frac{1}{4}\alpha x^4 \right\rangle_0 = \frac{3\alpha}{4K^2}(k_BT)^2.$$

That is,

$$A \le k_BT\log\left[\frac{\hbar\sqrt{K/m}}{k_BT}\right] + \frac{1}{2}(k/K-1)k_BT + \frac{3\alpha}{4K^2}(k_BT)^2.$$

Minimizing the right hand side wrt $K$, we obtain

$$\frac{1}{2K}-\frac{k}{2K^2}-\frac{3\alpha}{8K^3}k_BT=0.$$

If $\alpha=0$ clearly we get the right answer $K=k$. Solving this, we obtain

$$K=\frac{1}{2}\left(k+\sqrt{k^2+3\alpha k_BT}\right).$$

**2.23** [Gibbs-Bogoliubov's inequality (quantum case)][18]
Gibbs-Bogoliubov's inequality

$$A\le A_0+\langle H-H_0\rangle_0 \tag{6.2.119}$$

holds in quantum statistical mechanics as well.
(1) Demonstrate Peierls' inequaltiy:

$$Tre^{-\beta H}\ge\sum_i e^{-\langle i|H|i\rangle}, \tag{6.2.120}$$

where $\{|i\rangle\}$ is an arbitrary orthonormal basis.
(2) Let $\{|i\rangle\}$ be the orthonormal basis consisting of the eigenstates fo $H_0$. Then,

$$e^{-\beta A}\ge\sum_i e^{-\beta\langle i|H|i\rangle}=e^{-\beta A_0}\sum_i e^{\beta(A_0-\langle i|H_0|i\rangle)}e^{-\beta\langle i|(H-H_0)|i\rangle}. \tag{6.2.121}$$

Show Gibbs-Bogoliubov's inequality with the aid of Jensen's inequality.

**2.24** (1) For any density operator $P$

$$A\le TrP(H+k_BT\log P), \tag{6.2.122}$$

where $A$ is the free energy for the system whose hamiltonian is $H$.
(2) Suppose $P$ is the canonical density operator $P=e^{\beta(A_0-H_0)}$ for a system with the Hamiltonian $H_0$. Show that the above inequality is jsut Gibbs-Bogoliubov's inequality.

**2.25** [Convexity of free energy] (Ruellle)

$$A[\sum\lambda_iH_i]\ge\sum A[\lambda_iH_i]. \tag{6.2.123}$$

---

[18] M D Girardeau and R M Mazo, "Variational methods in statistical mechanics,"

**Soluton**
Hölder + Peierls proves this.

**2.26** [Thermodynamic perturbation theory]
Suppose the system Hamiltonian is given as $H = H_0 + \epsilon H_1$, where $\varepsilon$ is a (small) constant. Demnonstratet the following expansion formula:

$$A = A_0 + \varepsilon\langle H_1\rangle_0 - \frac{1}{2}\beta\varepsilon^2\langle(H_1 - \langle H_1\rangle_0)^2\rangle_0 + \cdots, \tag{6.2.124}$$

where $A$ is the free energy of the system, $A_0$ is the free energy in case $H_1 = 0$, and $\langle\ \rangle_0$ is the expectation with respect to the canonical distribution with the Hamiltonian $H_0$.

**2.27** [Jarzynski's equality].[19]
A single stranded DNA with a certain binding protein is stretched slowly until the protein dissociates from the DNA. Then, the length of the DNA is returned slowly to the rather relaxed state where the binding of the molecule does not affect the DNA tension. The work $W$ dissipated during the cycle is measured at 300K and the experimental results were as follows:

| $W$ in pNnm | number of times | $\beta W$ | $e^{-\beta W}$ |
|---|---|---|---|
| 78-82 | 4 | 19.3 | $4.04 \times 10^{-9}$ |
| 83-87 | 15 | 20.5 | $1.21 \times 10^{-9}$ |
| 88-92 | 7 | 21.74 | $3.62 \times 10^{-10}$ |
| 93-97 | 4 | 22.94 | $1.082 \times 10^{-10}$ |
| 98-102 | 1 | 24.15 | $3.23 \times 10^{-11}$ |

What is the best estimate of the (Gibbs) free energy change due to binding of the protein in the relaxed state of the single stranded DNA? How is your estimate different from the simple average $\langle W\rangle$?

**Solution**
Notice that $k_BT = 4.14$pNnm. $e^{-\beta W}$ is written in the above table. Thus,

$$\sum e^{-\beta W} = 373.1 \times 10^{-10} \quad \Rightarrow \quad \left\langle e^{-\beta W}\right\rangle = 1.2 \times 10^{-9} \tag{6.2.125}$$

That is, our estimate of $\Delta A$ is 85.0 pNnm. If we directly average the result, we obtain 87.4 pNnm. Of course, we have 'confirmed' the second law $\langle W\rangle \geq \Delta A$.

Although we wrote $A$ in the above, its definition is complicated.

---

[19] Inspired by Rustem Khafizov and Yan Chemla's experiment on SSB. The numbers are only fictitious. although the magnitudes are realistic.

**2.28** [Fluctuation and spring constant][20]
Inside the $F_1$ATPase is a rotor $\gamma$ to which a long actin filament (it is a straight stiff bar of length 30 nm) is perpendicularly attached. Thus, the filament swings back and forth when the ATPase is waiting for an ATP molecule.
(1) The root mean square angle fluctuation of the stiff filament was 30 degrees at 290K. If the temperature is raised by 10%, by what percentage will the angular fluctuation change? Assume that the molecular structure is not affected by this temperature change.
(2) What is the torsional spring constant of this rotor captured by the surrounding ring?
(3) Now, adding an appropriate polymers to the ambient solution, the effective viscosity of the solution is doubled. What is the mean square angle fluctuation of the filament? You may assume that the polymers do not affect the ATPase itself.

**Solution**
(1) Suppose $\theta$ is the angular deviation around the equilibrium direction. Then, the torsional spring constant $k$ reads

$$\tau = k\theta, \tag{6.2.126}$$

where $\tau$ is the torsion. Since, $k^{-1}$ is the 'susceptibility' of $\theta$ against $\tau$, the fluctuation-response relation tells us

$$k^{-1} = \left.\frac{\partial \theta}{\partial \tau}\right|_T = \beta\langle(\delta\theta)^2\rangle. \tag{6.2.127}$$

That is,

$$\langle(\delta\theta)^2\rangle = k_BT/k. \tag{6.2.128}$$

Since we may assume $k$ does not depend on $T$, the fluctuation should change by about 5%.
(2) You must measure the angle in radians.

$$k = 1.382 \times 10^{-23} \times 290/(\pi/6)^2 = 1.46 \times 10^{-20} \tag{6.2.129}$$

The unit is J/rad. Is it reasonable? It is about 15 pNnm/rad, a reasonable value.
(3) No change. The formula does not depend on the viscosity, so the amplitude of the fluctuation never changes. This is true however gooey the solution is. It is true that the oscillation becomes slow, but then small fluctuations can be accumulated to a size as large as when the viscosity is very low.

---

[20] If you wish to see the structure of the ATPase or to learn about how you can be alive, see K. Kinosita, K. Adachi, and H. Itoh, "Rotation of $F_1$ATPase: how an ATP-driven molecular machine may work," Ann. Rev. Biophys. Biomol. Struct., **33**, 245 (2005).

**2.29** [Thermodynamic fluctuations]
(1) Suppose $X$ and $y$ are nonconjugate pair with respect to energy, $X$ extensive and $y$ intensive. Prove that $\langle \delta X \delta y \rangle = 0$.
(2) Let $X$ and $x$ be a conjugate pair (wrt energy). Show $\langle \delta X \delta x \rangle = k_B T$.
(3) Express $\langle \delta \mu^2 \rangle$ in terms of a single thermodynamic derivative. The system is assumed to be described in terms of $S, V, N$ (or their conjugate variables).
(4) Show with the aid of grand partition function that

$$k_B T^2 \left.\frac{\partial E}{\partial T}\right|_{\mu,V} = \langle \delta E^2 \rangle - \mu \langle \delta E \delta N \rangle.$$

(5) Let $X$ be an extensive quantity. What can you conclude about $\langle \delta S \delta X \rangle$? The result is pedagogically suggestive, because entropy fluctuation means spatially local heat transport: that is, local temperature change.

**Solution**
(1) Recall that you can choose any combination of variables as independent variables as long as one variable is chosen from each conjugate pair $\{X, x\}$. We know $\langle \delta X_i \delta X_j \rangle$, so in this case, we should use the all extensive independent variable set.

$$\langle \delta X \delta y \rangle = \left\langle \delta X \sum_Y \frac{\partial y}{\partial \sum_Y} Y \delta Y \right\rangle = k_B T \sum_Y \frac{\partial Y}{\partial x} \frac{\partial y}{\partial Y} = k_B T \frac{\partial y}{\partial x} = 0.$$

In more detail for those who are skeptic:

$$dx = \sum_j \left.\frac{\partial x}{\partial X_j}\right|_{X_1 \cdots \check{X}_j \cdots X_n} dX_j, \tag{6.2.130}$$

where $\check{X}_j$ implies to remove the variable under the check mark. Therefore,

$$\left.\frac{\partial x_i}{\partial x_k}\right|_{x_1 \cdots \check{x}_k \cdots x_n} = \sum_j \left.\frac{\partial x_i}{\partial X_j}\right|_{X_1 \cdots \check{X}_j \cdots X_n} \left.\frac{\partial X_j}{\partial x_k}\right|_{x_1 \cdots \check{x}_k \cdots x_n}. \tag{6.2.131}$$

We put (6.2.130) into $\langle \delta X \delta y \rangle$ (regarding $X$ is a representative of $\{X_j\}$ and $y$ that of $\{x_k\}$ (the derivatives are mere constants, so you can take them out of the average symbol). Now, (6.2.131) tells you what you want.

(2)

$$\langle \delta X \delta x \rangle = \left\langle \delta X \frac{\partial x}{\partial Y} \delta Y \right\rangle = k_B T \sum_Y \frac{\partial Y}{\partial x} \frac{\partial x}{\partial Y} = k_B T.$$

(3) Taking into account of (1) above, we should choose $\mu, S, V$ as independent variables.

$$\delta^2 S = -\frac{1}{2T}(\delta N\delta\mu + \cdots) = -\frac{1}{2T}\left.\frac{\partial N}{\partial \mu}\right|_{S,V}\delta\mu^2 + \cdots.$$

Therefore,

$$P(\delta\mu\cdots) \propto \exp\left\{-\frac{1}{2k_BT}\left(\left.\frac{\partial N}{\partial \mu}\right|_{S,V}\delta\mu^2 + \cdots\right)\right\}.$$

That is,

$$\langle\delta\mu^2\rangle = k_BT\left.\frac{\partial \mu}{\partial N}\right|_{S,V}.$$

(4) Since

$$\Xi = \sum_{N=0}^{\infty} Z_N(T)e^{\beta\mu N} = \sum_{N=0}^{\infty}\int dE\, w_{E,N}e^{-\beta E+\beta\mu N},$$

$$\langle E\rangle = \frac{1}{\Xi}\sum_{N=0}^{\infty}\int dE\, w_{E,N}Ee^{-\beta E+\beta\mu N}.$$

Therefore,

$$\begin{aligned}\frac{d\langle E\rangle}{d\beta} &= \frac{1}{\Xi}\sum_{N=0}^{\infty}\int dEW_{E,N}E(\mu N - E)e^{-\beta E+\beta\mu N}\\ &\quad -\frac{1}{\Xi^2}\sum_{N=0}^{\infty}\int dE\, w_{E,N}Ee^{-\beta E+\beta\mu N}\sum_{N=0}^{\infty}\int dE\, w_{E,N}(\mu N - E)e^{-\beta E+\beta\mu N},\\ &= \langle\delta E(\mu\delta N - \delta E)\rangle\end{aligned}$$

(5)

$$\langle\delta S\delta X\rangle = k_BT\left.\frac{\partial X}{\partial T}\right|_{x}.$$

That is, the temperature derivative is the cross correlation with entropy fluctuation. This is, although trivial, worth remembering.

**2.30** [Equilibrium fluctuation]
(1) Obtain $\langle\delta S\delta V\rangle$.**q**
(2) Obtain $\langle\delta P\delta T\rangle$.

**Solution**

(1) Perhaps, the cleverest way is to us the fluctuation-response relation. We immediately obtain

$$\langle \delta S \delta V \rangle = k_B T \left.\frac{\partial V}{\partial T}\right|_P . \tag{6.2.132}$$

There is no simple trick, if you wish to use thermodynamic fluctuation theory. Choose $S$ and $V$ as independent variables.

$$\frac{1}{2k_BT}[\delta S \delta T - \delta P \delta V] = \frac{1}{2k_BT}\left[\left.\frac{\partial T}{\partial S}\right|_V + 2\left.\frac{\partial T}{\partial V}\right|_S \delta V \delta S - \left.\frac{\partial P}{\partial V}\right|_S \delta V^2\right]. \tag{6.2.133}$$

Therefore, (with the aid of the formulas for 2 variate Gaussian distribution)

$$\langle \delta S \delta V \rangle = k_B T \frac{1}{\frac{\partial(T,P)}{\partial(S,V)}} \left.\frac{\partial T}{\partial V}\right|_S = -k_B T \left.\frac{\partial S}{\partial P}\right|_T = k_B T \left.\frac{\partial V}{\partial T}\right|_P . \tag{6.2.134}$$

(2) There is no simple trick. Choose $T$ and $P$ as independent variables.

$$\frac{1}{2k_BT}[\delta S \delta T - \delta P \delta V] = \frac{1}{2k_BT}\left[\left.\frac{\partial S}{\partial T}\right|_P \delta T^2 + 2\left.\frac{\partial S}{\partial P}\right|_T \delta T \delta P - \left.\frac{\partial V}{\partial P}\right|_T \delta P^2\right]. \tag{6.2.135}$$

Therefore, (with the aid of the formulas for 2 variate Gaussian distribution)

$$\langle \delta T \delta P \rangle = k_B T \frac{1}{\left.\frac{\partial V}{\partial P}\right|_T \left.\frac{\partial S}{\partial T}\right|_P + \left.\frac{\partial S}{\partial P}\right|_T^2} \left.\frac{\partial S}{\partial P}\right|_T . \tag{6.2.136}$$

This is OK as an answer, but we can go further, if we realize

$$\langle \delta T \delta P \rangle = k_B T \frac{1}{\frac{\partial(V,S)}{\partial(P,T)}} \left.\frac{\partial S}{\partial P}\right|_T = k_B T \frac{\frac{\partial(S,T)}{\partial(P,T)}}{\frac{\partial(V,S)}{\partial(P,T)}} = k_B T \frac{\partial(S,T)}{\partial(V,S)} = -k_B T \left.\frac{\partial T}{\partial V}\right|_S . \tag{6.2.137}$$

**2.31** [Fluctuation and Le Chatelier-Braun's principle]
(1) Show that

$$\langle \delta x \delta X \rangle^2 \le \langle \delta x^2 \rangle \langle \delta X^2 \rangle$$

where $x$ and $X$ are conjugate pair of thermodynamic variables (wrt energy).
(2) What is the relation between this inequality and the Le Chatelier-Braun principle?

**Solution**
(1) The easiest way is to use the following obvious inequality valid for any real $t$:

$$0 \le \langle (\delta X + t\delta x)^2 \rangle = \langle \delta X^2 \rangle + 2t \langle \delta x \delta X \rangle + t^2 \langle \delta x^2 \rangle \tag{6.2.138}$$

Since $\langle \delta x^2 \rangle \geq 0$, we have its discriminant to be negative:

$$\langle \delta x \delta X \rangle^2 - \langle \delta X^2 \rangle \langle \delta x^2 \rangle \leq 0. \tag{6.2.139}$$

(2) We know $\langle \delta x \delta X \rangle = k_B T$, and (use clever way of calculating fluctuations)

$$\langle \delta X^2 \rangle = k_B T \left. \frac{\partial x}{\partial X} \right|_y^{-1}, \ \langle \delta x^2 \rangle = k_B T \left. \frac{\partial X}{\partial x} \right|_Y^{-1}. \tag{6.2.140}$$

Therefore,

$$\left. \frac{\partial x}{\partial X} \right|_y \leq \left. \frac{\partial x}{\partial X} \right|_Y. \tag{6.2.141}$$

Thus we have learned that Le Chatelier-Braun principle is a realization of $|\cos\theta| \leq 1$ (or Cauchy-Schwarz inequality) just as Hesenberg's uncertainty relation.

**2.32** [Fluctuation of internal energy]
For a classical monatomic ideal gas consisting of $N$ atoms, compute the fluctuation of its internal energy (under constant $T$ and $P$). Or show

$$\langle (E - \langle E \rangle)^2 \rangle / \langle E \rangle^2 = 2/3N. \tag{6.2.142}$$

**Solution**
The Gibbs relation $dE = TdS - PdV$ implies

$$\langle \delta E^2 \rangle = T^2 \langle \delta S^2 \rangle - 2TP \langle \delta S \delta V \rangle + P^2 \langle \delta V^2 \rangle. \tag{6.2.143}$$

The volume fluctuation can be found as

$$\langle \delta V^2 \rangle = -k_B T \left. \frac{\partial V}{\partial P} \right|_T = k_B \frac{V}{P}. \tag{6.2.144}$$

The entropy fluctuation can be calculated with the aid of $S$ and $P$ as independent variables, we conclude

$$\delta T = \left. \frac{\partial T}{\partial S} \right|_P \delta S + \cdots, \tag{6.2.145}$$

so

$$\langle \delta S^2 \rangle = T \left. \frac{\partial S}{\partial T} \right|_P = k_B C_P = \frac{5}{2} k_B^2 N. \tag{6.2.146}$$

Therefore, we need a result we have already obtained:

$$\langle \delta S \delta V \rangle = k_B T \left. \frac{\partial V}{\partial T} \right|_P = k_B V. \tag{6.2.147}$$

Combining all the results, we obtain

$$\langle \delta E^2 \rangle = k_B T^2 C_P - 2P k_B T^2 \left.\frac{\partial V}{\partial T}\right|_P - P^2 k_B T \left.\frac{\partial V}{\partial P}\right|_T, \tag{6.2.148}$$

$$= k_B^2 T^2 \left(\frac{5}{2}N - 2N + N\right) = \frac{3}{2}(k_B T)^2 N. \tag{6.2.149}$$

We know $\langle E \rangle = 3k_B T N/2$, so we arrive at the desired result.

**2.33** [Stability and related topics, e.g., Le Chatelier-Braun]
(1) Suppose a phase transition from phase I to phase II occurs upon increasing the magnetic field in the $z$-direction. What can you say about the relation between the magnetisms of the phases?
(2) Suppose phase I is a low temperature phase and II a high temperature phase. The phase transition I → II is first order. What can you say about the sign of the latent heat $\Delta H$ of this phase transition?
(3) Which specific heat is larger, $C_B$ or $C_M$ (under constant magnetic field, and under constant magnetization, respectively)?
(4) Suppose there is a dielectric material between a parallel plate capacitor. The two plates of the capacitor may be short-circuited by a switch. What can you say about the relation between the heat capacity of the dielectric material under the open- and short-circuited conditions? Let $\varepsilon$ be its dielectric constant, that may or may not depend on temperature.
(5) Suppose there is a liquid that crystallizes upon heating. Discuss the latent heat for this transition.[21]

**Solution**
(1) The internal energy must be convex, so the susceptibility must be nonnegative, if $M$ is differentiable with respect to $B$. At the phase transition this is not usually the case, but still the convexity must hold, so $M$ must increase in the second phase.
(2) The argument is the same as above (we did this problem before!). Increasing $T$ must increase $S$, so $S$ is larger for II. Therefore, $\Delta H = T\Delta S > 0$ if we go from I to II. That is latent heat must be absorbed by the system.
(3) This can be answered with the aid of Braun's principle:

$$\left.\frac{\partial x}{\partial X}\right|_y < \left.\frac{\partial x}{\partial X}\right|_Y. \tag{6.2.150}$$

---

[21] Johari, et al., "Endothermic freezing on heating and exothermic melting on cooling," J. Chem. Phys., **123**, 051104 (2005): $\alpha$-cyclodextrin + water + 4-methylpyridine (molar ratio of 1:6:100). For this system a liquid's endothermic freezing on heating and the resulting crystal's exothermic melting on cooling occur. $C_p$ decreases on freezing and increases on melting. Melting on cooling takes longer than freezing on heating.

Thus,

$$\left.\frac{\partial T}{\partial S}\right|_B = \frac{T}{C_B} < \left.\frac{\partial T}{\partial S}\right|_M = \frac{T}{C_M}. \tag{6.2.151}$$

That is, $C_B > C_M$: under constant $B$ $M$ is reorganized to absorb more heat. This is a hint to the next problem.

(4) When short-circuited, the electric field across the dielectric material is maintained to be constant (actually, zero). When, the circuit is open, then the surface charge (if any) on the dielectric material is kept constant; electric flux $D$ is maintained. Therefore, thermodynamically we expect $C_E \geq C_D$. Now, $\varepsilon$ is given, so we know the relation between $D$ and $E$: $D = \varepsilon EV$ ($D$ is extensive but $E$ is not! $D = \varepsilon E$ is a relation for a unit volume! Here, we assume $V$ does not change). We should proceed a step further. ($E$ in this problem is, of course, not the internal energy).

$$\left.\frac{\partial S}{\partial T}\right|_E = \left.\frac{\partial S}{\partial T}\right|_D + \left.\frac{\partial S}{\partial D}\right|_T \left.\frac{\partial D}{\partial T}\right|_E. \tag{6.2.152}$$

With the aid of a Maxwell's relation we obtain

$$\left.\frac{\partial S}{\partial D}\right|_T = -\left.\frac{\partial E}{\partial T}\right|_D = \frac{D}{\varepsilon^2 V}\frac{d\varepsilon}{dT}. \tag{6.2.153}$$

Also

$$\left.\frac{\partial D}{\partial T}\right|_E = \frac{D}{\varepsilon}\frac{d\varepsilon}{dT}. \tag{6.2.154}$$

Therefore,

$$C_E = C_D + T\frac{ED}{\varepsilon^2}\left(\frac{d\varepsilon}{dT}\right)^2. \tag{6.2.155}$$

This tells us that if $\varepsilon$ does not depend on $T$, then there is no difference between the two specific heat. This should be intuitively obvious, because no 'reorganization' of the material is expected upon heating.

(5) The original paper contains the answer.

**2.34** [Chemical equilibrium constant[22]]

The reaction

$$A \underset{k_-}{\overset{k_+}{\rightleftarrows}} B. \tag{6.2.156}$$

---

[22] A. B. Adib, "Symmetry Relations in Chemical Kinetics Arising from Microscopic Reversibility," Phys. Rev. Lett., **96**, 028307 (2006).

may be described as follows, if $A$ and $B$ are sufficiently dilute:

$$\frac{d[A]}{dt} = k_-[B] - k_+[A] = -\frac{d[B]}{dt}. \tag{6.2.157}$$

For all $t > 0$ show that

$$\frac{[B]_F(t)}{[A]_B(t)} = K, \tag{6.2.158}$$

holds. Here, $F$ denotes the forward reaction starting with pure A, and $R$ denotes the reverse reaction starting with the same moles of B as A. That is, if these two reactions are started simultaneously, then the concentration ratio at time $t$ as (6.2.158) is time-independent and equal to the chemical equilibrium constant. [However, this cannot be a general relation, but holds only under ideal solution and reaction conditions.]

**Solution**

Since $[A] + [B] = C$ (constant), $[B]_F(t)$ obeys

$$\frac{d[B]_F(t)}{dt} = -(k_+ + k_-)[B]_F(t) + k_+C$$

with the initial condition $[B]_F(0) = 0$. Similarly,

$$\frac{d[A]_B(t)}{dt} = -(k_+ + k_-)[A]_B(t) + k_-C$$

with the initial condition $[A]_B(0) = 0$. Therefore,

$$[B]_F(t) = \frac{k_+}{k_- + k_+}C\left(1 - e^{-(k_-+k_+)t}\right),$$

$$[A]_B(t) = \frac{k_-}{k_- + k_+}C\left(1 - e^{-(k_-+k_+)t}\right).$$

These formulas tell us what we wish to have, because $K = k_+/k_- = [B]_{eq}/[A]_{eq}$.

## 6.3 Problems for Chapter 3

**3.1** [Fermions and bosons; the ultimate elementary problem]
There is a system with only three states with energies 0, $\varepsilon$ and $\varepsilon$ ($\varepsilon > 0$, i.e., excited states are degenerate). There are three identical particles without spin.
(1F) If the particles are fermions, write down the canonical partition function.
(2F) Find the probability of finding $N$ particles in the ground state.
(3F) Compute the average occupation number $N_0$ of the ground state.
(1-3B) Repeat the same problems assuming that the particles are bosons.
(4) Consider the high temperature limit. (UIUC Qual Spring00)

**Solution**
(1F) Since all the one-particle states must be occupied, and there is only one microstate for the system:

$$Z = e^{-2\beta\varepsilon}. \tag{6.3.1}$$

(2F) Since the one-particle ground state is always occupied by a particle, $P(N) = \delta_{N,1}$.
(3F) 1; there is no fluctuation at all.
(1-3B) The microstates of the system may be classified according to the number of particles occupying the one-particle ground state: $n = 3$, 2, 1, or 0. They respectively correspond to the microstates with the total energy 0, $\varepsilon$, $2\varepsilon$ or $3\varepsilon$. The degeneracy of the macrostate designated by $n$ is $\binom{3-n+1}{1} = 4-n$. Therefore, the canonical partition function is given by

$$Z = 1 + 2e^{-\beta\varepsilon} + 3e^{-2\beta\varepsilon} + 4e^{-3\beta\varepsilon}. \tag{6.3.2}$$

The probability that the one-particle ground state is occupied by $n$ particles is given by

$$P(n) = \frac{1}{Z}(4-n)e^{-(3-n)\beta\varepsilon}. \tag{6.3.3}$$

The expectation value $\langle n \rangle$ is

$$\langle n \rangle = \frac{3 + 4e^{-\beta\varepsilon} + 3e^{-2\beta\varepsilon}}{1 + 2e^{-\beta\varepsilon} + 3e^{-2\beta\varepsilon} + 4e^{-3\beta\varepsilon}}. \tag{6.3.4}$$

(4) For the fermion case nothing changes even at high temperatures. For bosons in the $\beta \to 0$ limit, all 10 microstates are equally probable:

$$P(n) = (4-n)/10, \quad \langle n \rangle = 1. \tag{6.3.5}$$

**3.2** [Elementary problem for boson systems]
There are 100 identical spinless bosons whose $s$-th one-particle state has an energy $E_s = s\varepsilon$ ($s \in \mathbf{N}$) and is described by a wave function $\phi_n(r)$ (normalized). These particles do not interact.
(1) How many microstates with the energy $4\varepsilon$ does the system have?
(2) When the system is in equilibrium with the particle reservoir (chemostat) of temperature $T$ and chemical potential $\mu$, on the average 99 particles occupy the one-particle ground state ($s = 0$), and one particle occupies the one-particle first excited state ($s = 1$). The other one-particle states are negligibly occupied, Find $\mu$ and $\beta = 1/k_BT$ in terms of $\varepsilon$.

**Solution**
(1) This is a problem of partitioning an integer. The microstates with the total energy $4\varepsilon$ are

$$\begin{aligned} 4 &= 1+1+1+1, && (6.3.6)\\ &= 1+1+2, && (6.3.7)\\ &= 2+2, && (6.3.8)\\ &= 1+3, && (6.3.9)\\ &= 4. && (6.3.10)\end{aligned}$$

That is, there are 5 distinct microstates.
(2) Since

$$\begin{aligned} \langle n_0\rangle &= \frac{1}{e^{-\beta\mu}-1} = 99, && (6.3.11)\\ \langle n_1\rangle &= \frac{1}{e^{\beta(\varepsilon-\mu)}-1} = 1, && (6.3.12)\end{aligned}$$

we have

$$\begin{aligned} -\beta\mu &= \log(100/99) = 1.005\times 10^{-2}, && (6.3.13)\\ \beta(\varepsilon-\mu) &= \log 2 = 0.693. && (6.3.14)\end{aligned}$$

Hence, $\beta = 0.692/\varepsilon$ and $\mu = -0.01445\varepsilon$. Clearly recognize that $\mu$ is negative (does not exceed the ground state energy)!

**3.3** [Basic problem for quantum ideal gas: isothermal compression]
In a cylinder with a piston are $N$ identical particles. The temperature is maintained constant.
The fermion case:

(1) Suppose the system is maintained at $T = 0$, and the volume has been reversibly halved. What is the relation between the initial energy $e_i$ per particle and the final energy $e_f$ per particle?
(2) In the process described in (1) what is the ratio $P_f/P_i$, where $P_i$ is the initial pressure and $P_f$ the final pressure.
(3) Now, suppose the system is maintained at a positive temperature $T$. As in (1) we halve the volume of the system reversibly . How does the ratio $e_f/e_i$ change as a function of $T$? You may assume $T$ is sufficiently close to $T = 0$.
The boson case:
(4) Suppose the density of the condensate is positive at the initial temperature. After the volume is halved reversibly does the density of the condensate remain positive?
(5) Suppose $T = 0$ when the volume is reversibly halved. Find the ratio $P_f/P_i$ , where $P_i$ is the initial pressure and $P_f$ the final pressure.

**Solution**
(1) Let us write $gD_t(\varepsilon) = aV\varepsilon^{1/2}$ with a numerical constant $a$. We know the following relation:

$$N = \frac{2}{3}gD_t(\mu(0))\mu(0) = \frac{2}{3}aV\mu(0)^{3/2}, \tag{6.3.15}$$

$$E = \frac{2}{5}gD_t(\mu(0))\mu(0)^2 = \frac{2}{5}aV\mu(0)^{5/2}. \tag{6.3.16}$$

From these formulas we get the relation

$$E/N = \frac{3}{5}\mu(0) \tag{6.3.17}$$

we already know. Since $N$ is constant, $\mu(0) \propto V^{-2/3}$, so

$$e_f/e_i = (V/2)^{-2/3}/V^{-2/3} = 2^{2/3}. \tag{6.3.18}$$

(2) We use the universal relation $P \propto E/V$ for any ideal gas

$$P_f/P_i = 2(e_f/e_i) = 2^{5/3}, \tag{6.3.19}$$

where the factor 2 comes from the volume ratio $V_f/V_i$.
(3) In contrast to the $T = 0$ case, the particles need not be pushed up with the energy levels. Consequently, the increase ratio of the energy is expected to decrease with $T$.
(4) If the volume is decreased, the energy level spacings widen. Therefore, more particles fall to the ground state. That is, $N_0$ should increase. Quantitatively, we have only to look at $N_1 = AVT^{3/2}$. Since $T$ is kept constant, $N_1$ halves.

(5) Since $E \propto VT^{5/2}$, $P = (2/3)(E/V)$ does not depend on the system volume. Hence, there is no pressure change: the ratio is unity. We can regard the ground state as a pressure buffer.

**3.4** [Basic problem for quantum ideal gas: adiabatic free expansion]
In a thermally isolated cylinder with a piston is an ideal gas, whose initial temperature is $T_i$. The piston is pulled suddenly to increase the volume by 10%.

The fermion case: Suppose the ideal gas is fermionic.
(1F) Find the final pressure $P_f$ in terms of $P_i$, the initial pressure.
(2F) Which is correct, $T_i < T_f$, $T_i = T_f$ or $T_i > T_f$?
(3F) Suppose the initial temperature is $T = 0$. Express the final temperature $T_f$ approximately in terms of the Fermi temperature $T_F$ before the expansion.

The boson case: Suppose the ideal gas is bosonic.
(1B) Find the final pressure $P_f$ in terms of $P_i$, the initial pressure.
(2B) Suppose the initial temperature is sufficiently low and the condensate does not disappear by expansion. What is the final temperature $T_f$?
(3B) Suppose the initial temperature is less than $T_c$. After expansion, the final temperature becomes exactly $T_c$ (for the expanded system). Find the initial temperature $T_i$ in terms of the $T_c$ *before* expansion.

**Solution**
(1F) The internal energy $E$ does not change, because the system is adiabatic and free expansion does not do work. Therefore,

$$PV = \frac{2}{3}E \tag{6.3.20}$$

implies

$$P_iV = P_f(1.1V). \tag{6.3.21}$$

Hence, $P_f = 0.91P_i$.
(2F) Expansion makes packing energy level denser, so to keep the total energy the only way is to occupy more excited states. Therefore, $T_i < T_f$.

This relation cannot be obtained purely thermodynamically. We have

$$\left.\frac{\partial T}{\partial V}\right|_E = \frac{\partial(T,E)}{\partial(V,T)}\frac{\partial(V,T)}{\partial(V,E)} = -\frac{1}{C_V}\left.\frac{\partial E}{\partial V}\right|_T, \tag{6.3.22}$$

(this is thermodynamics) and

$$\left.\frac{\partial E}{\partial V}\right|_T = \frac{3}{2}\left[\left.\frac{\partial P}{\partial V}\right|_T V + P\right] \tag{6.3.23}$$

(this is no more a thermodynamic relation, since we have used (**??**)). To proceed further, we need the equation of state. Let us first consider the classical case $PV = Nk_BT$:

$$V\left.\frac{\partial P}{\partial V}\right|_T = -\frac{Nk_BT}{V} = -P \tag{6.3.24}$$

That is, $T_i = T_f$ for the classical case. Compared with this case, fermionic ideal gas should be harder to compress, so $(\partial E/\partial V)_T < 0$, which implies $(\partial T/\partial V)_E > 0$. We could use an explicit energy formula as you can see in the following discussion.
(3F)

$$E = \frac{3}{5}N\mu(0) + \zeta(2)\frac{3}{2\mu(0)}N(k_BT)^2 + \cdots, \tag{6.3.25}$$

where

$$\mu(0) = \frac{h^2}{2m}\left(\frac{3N/V}{4\pi}\right)^{2/3}. \tag{6.3.26}$$

If $V$ is increased, the Fermi energy Fermi $\mu(0)$ goes down, so the first term of the above formula decreases. Therefore, to keep $E$ constant, we must increase the second term. That is, $T$ goes up. The Fermi temperature after expansion of the volume is Ferm $(1/1.1)^{2/3}T_F = 0.938T_F$, so approximately

$$T^2 \simeq \frac{4}{3\pi^2}0.062(\mu(0)/k_B)^2 = \frac{4}{3\pi^2}0.062T_F^2 = 0.00838T_F^2, \tag{6.3.27}$$

or $T = 0.092T_F$.
(1B) This is the same as (1F).
(2B) This is due to $E \propto VT^{5/2}$. $T_f = 1.1^{-0.4}T_i \simeq 0.962T_i$, so $T_f < T_i$. In this case as well the expansion makes the energy level packing denser, so consequently excitations become easier and the amount of condensate should decrease. However, this does not imply the increase of temperature, because the total energy can be maintained constant by occupying lower energy states with more particles.
(3B) (3.4.23) implies that the $T_c$ before expansion is (here, $m$ is the mass of the particle)

$$T_c = \frac{h^2}{2\pi k_Bm}\left(\frac{n}{\zeta(3/2)}\right)^{2/3} \tag{6.3.28}$$

Therefore, $1.1(T_c^*)^{5/2} = T_i^{5/2}$, where $T_c^*$ is the critical temperature after expansion. Hence,

$$1.1^{-2/5}T_i = T_c^* = \frac{h^2}{2\pi k_Bm}\left(\frac{n/1.1}{\zeta(3/2)}\right)^{2/3}, \tag{6.3.29}$$

where $n$ is the number density before expansion. In terms of $T_c$

$$1.1^{-2/5+2/3}T_i = T_c \;\Rightarrow T_i = 1.1^{-4/15}T_c = 0.975T_c \tag{6.3.30}$$

**3.5** [Basic problem for quantum ideal gas: adiabatic quasistatic expansion]
In a thermally isolated cylinder with a piston is an ideal gas, whose initial temperature is $T_i$ and initial pressure is $P_i$. The piston is pulled slowly to double the volume.

The fermion case: Suppose the ideal gas is fermionic.
(1F) Obtain the final pressure $P_f$ in terms of $P_i$.
(2F) What is the final temperature $T_f$, if $T_i = 0$?
(3F) More generally, obtain $T_f$ in terms of $T_i$.

The boson case: Suppose the ideal gas is fermionic.
(1B) Obtain the final pressure $P_f$ in terms of $P_i$.
(2B) Obtain $T_f$ in terms of $T_i$, assuming that the condensate does not disappear.
(4B) Let $N_{0i}$ be the initial number of particles in the condensate. Does the final number of particles $N_{0f}$ in the condensate increase or decrease?

**Solution**

(1F) In this case, entropy does not change, so $dE = -PdV$. For any ideal gas $PV = (2/3)E$, so

$$dE = -\frac{2}{3}\frac{E}{V}dV. \tag{6.3.31}$$

This implies that $EV^{2/3}$ is constant. As can be seen from this derivation, the ration is independent of statistics. Since $E_iV^{2/3} = E_f(2V)^{2/3}$, we get $E_f = 2^{-2/3}E_i$. That is,

$$P_f(2V) = \frac{2}{3}E_f = \frac{2}{3}2^{-2/3}E_i = 2^{-2/3}P_iV, \tag{6.3.32}$$

or

$$P_f = \frac{1}{2^{5/3}}P_i. \tag{6.3.33}$$

(2F) We can expect $T_f = 0$. Indeed, at $T = 0$

$$E = \int_0^{\mu(0)} d\varepsilon D_t(\varepsilon)\varepsilon = \frac{3}{4}\mu(0)N \propto N^{5/3}V^{-2/3}. \tag{6.3.34}$$

That is, $EV^{2/3}$ is kept constant (adiabatic and isothermal processes can agree only at $T = 0$). This result is an example of the general rule that the gas temperature never goes up through quasistatic expansion.

(3F) This can be solved with the aid of (6.3.25).

(1B) Exactly the same as (1F).
(2B) Since BCE occurs, $\mu = 0$ and

$$E = \int_0^\infty d\varepsilon D_t(\varepsilon) \frac{\varepsilon}{e^{\beta\varepsilon} - 1} \propto VT^{5/2}. \tag{6.3.35}$$

We know $EV^{2/3}$ is maintained constant irrespective of statistics, $V^{5/3}T^{5/2}$ is invariant. That is, $V^{2/3}T$ is constant. Hence, $T_f = 2^{-2/3}T_i$; the system temperature goes down.

Notice, however, that if we admit that the pressure increases with temperature, thermodynamics can tell this:

$$\left.\frac{\partial T}{\partial V}\right|_S = -\frac{T}{C_V}\left.\frac{\partial S}{\partial V}\right|_T = -\frac{T}{C_V}\left.\frac{\partial P}{\partial T}\right|_V < 0. \tag{6.3.36}$$

(4B) Since the process we are interested in is quasistatic and adiabatic, the average occupation number of the one-particle ground state should not change. If you realize this no calculation is needed, but if you wish to confirm this by computation, use $N_{0i} = N - N_{1i}$ and $N_{1i} = cVT_i^{3/2}$:

$$N_{1f} = c(2V)T_f^{3/2} = cVT_i^{3/2} = N_{1i}. \tag{6.3.37}$$

Hence, $N_0$ cannot change.

**3.6** [Basic problem for quantum ideal gas: compression under constant internal energy]
In a cylinder with a piston is an ideal gas consisting of $N$ particles, whose initial temperature is $T_i$. The piston is pushed in slowly to halve the volume while removing thermal energy appropriately to keep the internal energy constant. Let $T_f$ be the final temperature.

I.The case of spinless bosons: assume that there is a Bose-Einstein condensate initially.
(1) Find the number of particles $N_0$ in the condensate before compression. You may use the critical temperature $T_c$.
(2) Which is true, $T_f < T_i$, $T_f = T_i$ or $T_f > T_i$?
(3) Does the number of particles in the condensate increase or decrease?

II. The case of spin 1/2 fermions.
(4) Find the final pressure $P_f$.
(5) Is there a minimum temperature ($> 0$) below which this process becomes impossible?

(6) Which is true, $T_f < T_i$, $T_f = T_i$ or $T_f > T_i$?

**Solution**

(1)

$$N_0 = N(1 - (T/T_c)^{3/2}), \tag{6.3.38}$$

where $T_c \propto n^{2/3}$.

(2) Below $T_c$ we know $E \propto VT^{5/2}$. Therefore,

$$VT_i^{5/2} = (V/2)T_f^{5/2}. \tag{6.3.39}$$

That is, $T_f = 2^{2/5}T_i$ or $T_f > T_i$.

(3) Since $T_c \propto n^{2/3}$, $T_{cf} = T_{ci}2^{2/3}$ and $T_f = 2^{2/5}T_i$ hold. Therefore,

$$T_f/T_{cf} = 2^{2/5-2/3}T_i/T_{ci} < T_i/T_{ci}. \tag{6.3.40}$$

Consequently, the ration in (6.3.38) decreases and $N_0$ increases. This is also understandable from the widening of the energy level spacings.

(4) Since

$$P_iV = 2E/3 = P_f(V/2), \tag{6.3.41}$$

we get $P_f = 2P_i = 4E/3V$.

(5) At $T = 0$, we know $E \propto n^{2/3}$, so $E$ increases if the system is compressed; this should be intuitively obvious because the level spacings increase. We cannot cool the system further if $T$ is very low. Therefore, quasiequilibrium constant energy process becomes impossible at some low but positive temperature.

(6) We can generally write

$$\left.\frac{\partial E}{\partial V}\right|_T = T\left.\frac{\partial P}{\partial T}\right|_V - P = \frac{2}{3V}\left[T\left.\frac{\partial E}{\partial T}\right|_V - E\right]. \tag{6.3.42}$$

For ideal Fermi gases the graph of $E(T)$ (Fig. 3.3.3) implies

$$\frac{E}{T} > \left.\frac{\partial E}{\partial T}\right|_V. \tag{6.3.43}$$

The difference converges to zero in the high temperature limit; the inequality is not due to thermodynamics. This inequality combined with (6.3.42) implies that $(\partial E/\partial V)_T < 0$. That is, if $T$ were kept constant and the system volume decreased, then $E$ would increase. Therefore, to maintain $E$, heat would have to be discarded. Thus, the final temperature must be smaller: $T_i > T_f$. This conclusion can also be obtained by noting that the energy level spacing increases upon compression.

**3.7** [Qualitative properties of quantum ideal gases]
Assume the particles do not interact. Answer the following qualitative questions and give your justification for your answers. All the processes are quasistatic.

The boson case: there are $N$ bosons in a volume $V$.
(1B) The volume is increased under constant energy. Does the temperature decrease?
(2B) The volume is increased under constant entropy. Does the temperature decrease?
(3B) Can we decrease the volume while keeping the internal energy?

The fermion case: there are $N$ fermions in a volume $V$.
(1F) The volume is increased under constant energy. Does the temperature decrease?
(2F) The volume is increased under constant entropy. Does the temperature decrease?
(3F) Can we decrease the volume while keeping the internal energy?

**Solution**
(1B) Below $T_c$ we can write explicitly as $E \propto VT^{5/2}$ ($5/2 = d/\alpha + 1$), so we immediately see that $T$ decreases. If the volume is increased, the energy level spacing decreases, so excitations become easier (consequently $T_c$ goes down), so the number of particles occupying the one-particle ground state decreases. If you wish to keep the system energy despite this, you have to decrease the system temperature. What could happen above $T_c$ is subtle, as can be seen from the behavior of $C_V$. If the temperature is sufficiently high, then the system is close to a classical ideal gas, so the temperature dependence diminishes.
(2B) If $S$ is kept, the particles must follow the behavior of the energy levels. The level spacings decrease, so this is possible only by decreasing the temperature.
(3B) This is possible, if heat is supplied appropriately to warm up the system.
(1F) The energy level spacings decrease, so the total energy cannot be maintained without increasingly occupying the excited states. Hence, the temperature goes up.
(2F) To keep $S$, the shape of the 'cliff' of the fermi distribution, but since the Fermi energy goes down, this is possible only through cooling the system.
(3F) This is generally impossible.

**3.8** [Conversion of fermion into bosons]
There is an ideal Fermi gas with the total energy 10 eV in an adiabatic container. The fermion particles are actually metastable and turn into bosons without adding any energy. Assume that the conversion is done quasistatically and adiabatically. Does the container explode? [UIUC qual]

**Solution**

Irrespective of statistics $PV = 2E/3$. Since $E$ and $V$ are constant, the pressure does not change.

We know, if $N$ and $T$ are identical $P_{FD} > P_{MB} > P_{BE}$ (→(3.1.17)). Since $E$ is an increasing function of $T$, the pressure of the gas is an increasing function of $T$ (however, $(\partial P/\partial T)_V > 0$ is NOT a thermodynamic inequality; think of counter examples). Therefore, we must conclude $T_{FD} < T_{MB} < T_{BE}$. This suggests that we may handle the boson system as a classical ideal gas system; we can easily guess $T \sim T_F$. Indeed, we can estimate the system temperature after conversion as

$$E = (3/5)\mu_0 n = (3/2)nk_BT \Rightarrow T = 2T_F/5. \tag{6.3.44}$$

This is an extreme high temperature, so the container melts away, and there is an explosion.

**3.9** [Equation of state of ideal gases]
We know the relation between $PV$ and the internal energy does not depend on particle statistics.
(1) Is this still true for ideal gas mixtures?
(2) Compute $PV/E$ in $d$-space (this is already mentioned in the text).

**Solution**
(1) You may use the law of partial pressure. Let $P_i$ be the partial pressure due to chemical species $i$. If its internal energy is $E_i$, $P_iV = (2/3)E_i$ holds for all $i$. Since the internal energy is additive $PV = (2/3)E$ must also hold for the ideal gas mixtures.
(2) You have only to trace the proof of $PV = 2E/3$. See up to (3.1.24). Let $D_t(\varepsilon)$ be the density of translational states in $d$-space. The key element of the derivation of (3.1.24) is the relation between $d(\varepsilon D(\varepsilon))/d\varepsilon$ and $D(\varepsilon)$. In $d$-space we can write $D(\varepsilon) = A\varepsilon^{d/2-1}$ with an appropriate numerical factor $A$, so

$$\frac{d}{d\varepsilon}\{\varepsilon D(\varepsilon)\} = \frac{d}{2}D(\varepsilon). \tag{6.3.45}$$

Hence, we get $PV/E = 2/d$.

**3.10** [Effective interaction due to statistics]
Fig. 3.1.1 illustrates how we can intuitively understand the effective interactions between particles: compared with classical particles, between bosons there is an effective attraction, and between fermions there is an effective repulsion. Let us make this understanding slightly quantitative. Here, we proceed step by small step, reviewing elementary quantum mechanics.

We wish to consider a two-particle system in terms of canonical ensemble theory.

The system Hamiltonian reads

$$H = \frac{\boldsymbol{p}_1^2}{2m} + \frac{\boldsymbol{p}_2^2}{2m}, \tag{6.3.46}$$

and the canonical partition function is

$$Z = Tre^{-\beta H}, \tag{6.3.47}$$

where the trace is with respect to the microstates specified by two momenta $|\boldsymbol{p}, \boldsymbol{p}'\rangle$. To compute this trace semi-classically, we introduce a single-particle momentum state $|\boldsymbol{p}\rangle$.
(1) Express $|\boldsymbol{p}, \boldsymbol{p}'\rangle$ both for the boson and fermion cases in terms of single particle kets $|\boldsymbol{p}\rangle$. You may regard two momenta are distinct, but the obtained states must be properly normalized.
(2) Assuming that the system is in a sufficiently big box of volume $V$, find the position representation $\langle \boldsymbol{r}|\boldsymbol{p}\rangle$ (i.e., the wave function) of the momentum ket $|\boldsymbol{p}\rangle$.
(3) Let $\boldsymbol{r}_i$ be the position vector of the $i$-th particle. Find the position representation of $|\boldsymbol{p}, \boldsymbol{p}'\rangle$. [This is of course virtually the same question as (1).]
(4) For an $N$-particle system in the semi-classical limit, the calculation of trace in $Z$ may be performed as follows:

$$Tr \;\to\; \frac{1}{N!}\int_{V^N} d\{\boldsymbol{r}_k\} \prod_{k=1}^{N} \langle \boldsymbol{r}_k| \cdots \prod_{k=1}^{N} |\boldsymbol{r}_k\rangle \tag{6.3.48}$$

$$= \frac{1}{N!}\int_{V^N} d\{\boldsymbol{r}_k\} \prod_{k=1}^{N} \langle \boldsymbol{r}_k| \left[\left(\sum_{\{\boldsymbol{p}_i\}} |\{\boldsymbol{p}_i\}\rangle\langle\{\boldsymbol{p}_i\}|\right) \cdots \left(\sum_{\{\boldsymbol{p}_i\}} |\{\boldsymbol{p}_i\}\rangle\langle\{\boldsymbol{p}_i\}|\right)\right] \prod_{k=1}^{N} |\boldsymbol{r}_k\rangle \tag{6.3.49}$$

If the volume is big enough, we should be able to replace the summation over momenta by integration over them. The replacement rule is

$$\sum_{\{\boldsymbol{p}_i\}} \to \frac{V^N}{h^{3N}} \int d\{\boldsymbol{p}_i\}. \tag{6.3.50}$$

Justify this for $N = 1$ in 1-space.
(5) Write $Z$ down using $h^{-3/2} e^{i\boldsymbol{r}\cdot\boldsymbol{p}/\hbar} = \langle \boldsymbol{r}_i|\boldsymbol{p}\rangle$. Beyond this point, let us simplify formulas by taking the $V \to \infty$ limit. You need not perform the integration.
(6) The outcome of (5) must have the following form:

$$\frac{1}{2h^6}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 d\boldsymbol{p} d\boldsymbol{p}' e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}[\cdots]. \tag{6.3.51}$$

Perform the integral in this expression and find $F$ in the following formula:

$$Z = \frac{1}{2h^6}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 d\boldsymbol{p} d\boldsymbol{p}'\, e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m} F \tag{6.3.52}$$

(7) $F$ may be interpreted as the Boltzmann factor coming from the effective interaction originating from particle statistics. Sketch the potential ($\times\beta$) of this effective interaction for bosons and fermions.

**Solution**

(1) The ket $|\boldsymbol{p}\rangle|\boldsymbol{p}'\rangle$ must be correctly symmetrized; $+$ is for bosons and $-$ for fermions:

$$|\boldsymbol{p},\boldsymbol{p}'\rangle = \frac{1}{\sqrt{2}}(|\boldsymbol{p}\rangle|\boldsymbol{p}'\rangle \pm |\boldsymbol{p}'\rangle|\boldsymbol{p}\rangle). \tag{6.3.53}$$

(2) $|\boldsymbol{p}\rangle$ describes a plane wave of wave vector $\boldsymbol{k} = \boldsymbol{p}/\hbar$:

$$\langle \boldsymbol{r}|\boldsymbol{p}\rangle \propto e^{i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}. \tag{6.3.54}$$

The normalization condition is

$$\delta_{\boldsymbol{p}\boldsymbol{p}'} = \frac{1}{\hbar}\int_V d^3\boldsymbol{r}\, \langle \boldsymbol{p}'|\boldsymbol{r}\rangle\langle \boldsymbol{r}|\boldsymbol{p}\rangle. \tag{6.3.55}$$

Therefore,

$$\langle \boldsymbol{r}|\boldsymbol{p}\rangle = \sqrt{\frac{\hbar}{V}} e^{i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}. \tag{6.3.56}$$

(3)

$$(\langle \boldsymbol{r}_1|\langle \boldsymbol{r}_2|)|\boldsymbol{p},\boldsymbol{p}'\rangle = \frac{1}{\sqrt{2}}(\langle \boldsymbol{r}_1|\boldsymbol{p}\rangle\langle \boldsymbol{r}_2|\boldsymbol{p}'\rangle \pm \langle \boldsymbol{r}_1|\boldsymbol{p}'\rangle\langle \boldsymbol{r}_2|\boldsymbol{p}\rangle). \tag{6.3.57}$$

(4) The left-hand side is the sum over all the states in the volume $V$. If we adopt a periodic boundary condition $k = (2\pi/L)n$ ($n \in \boldsymbol{Z}$).

$$\sum_{n=-\infty}^{\infty} n \simeq \int_{-\infty}^{\infty} dn = \frac{L}{2\pi}\int \frac{2\pi}{L} dn = \frac{L}{2\pi}\int dk = \frac{L}{h}\int dp. \tag{6.3.58}$$

The 3-dimensional version reads

$$\sum_{\boldsymbol{p}} \simeq \frac{V}{h^3}\int d\boldsymbol{p}. \tag{6.3.59}$$

If you do not like the periodic boundary condition, $k = (\pi/L)n$ ( $n \in \boldsymbol{N}$) and

$$\sum_{n=1}^{\infty} \simeq \int_1^{\infty} dn = \frac{L}{\pi}\int_1^{\infty} \frac{\pi}{L}dn = \frac{L}{\pi}\int_0^{\infty} dk = \frac{L}{h}\int_{-\infty}^{\infty} dp. \tag{6.3.60}$$

(5) Using the results of (2) and (3), we get (the overall factor 1/2 comes from $1/N!$ in the definition of trace)

$$Z = Tr\, e^{-\beta H} = \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \langle \boldsymbol{r}_1|\langle \boldsymbol{r}_2|e^{-\beta H}|\boldsymbol{r}_1\rangle|\boldsymbol{r}_2\rangle, \tag{6.3.61}$$

$$= \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \sum_{\boldsymbol{p},\boldsymbol{p}'} e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}|(\langle \boldsymbol{r}_1|\langle \boldsymbol{r}_2|)|\boldsymbol{p},\boldsymbol{p}'\rangle|^2, \tag{6.3.62}$$

$$= \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \sum_{\boldsymbol{p},\boldsymbol{p}'} e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}|(\langle \boldsymbol{r}_1|\langle \boldsymbol{r}_2|)|\boldsymbol{p},\boldsymbol{p}'\rangle|^2, \tag{6.3.63}$$

$$= \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \sum_{\boldsymbol{p},\boldsymbol{p}'} e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}\frac{1}{2}|\langle \boldsymbol{r}_1|\boldsymbol{p}\rangle\langle \boldsymbol{r}_2|\boldsymbol{p}'\rangle \pm \langle \boldsymbol{r}_1|\boldsymbol{p}'\rangle\langle \boldsymbol{r}_2|\boldsymbol{p}\rangle|^2. \tag{6.3.64}$$

If we write the matrix elements explicitly,

$$Z = \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \sum_{\boldsymbol{p},\boldsymbol{p}'} e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}\frac{1}{V^2}[1 \pm Re\exp(i(\boldsymbol{p}-\boldsymbol{p}')\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)/\hbar)], \tag{6.3.65}$$

$$= \frac{1}{2h^6}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \int d\boldsymbol{p} d\boldsymbol{p}' e^{-\beta(\boldsymbol{p}^2+\boldsymbol{p}'^2)/2m}[1 \pm Re\exp(i(\boldsymbol{p}-\boldsymbol{p}')\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)/\hbar)]. \tag{6.3.66}$$

(6) To obtain $F$ we compute

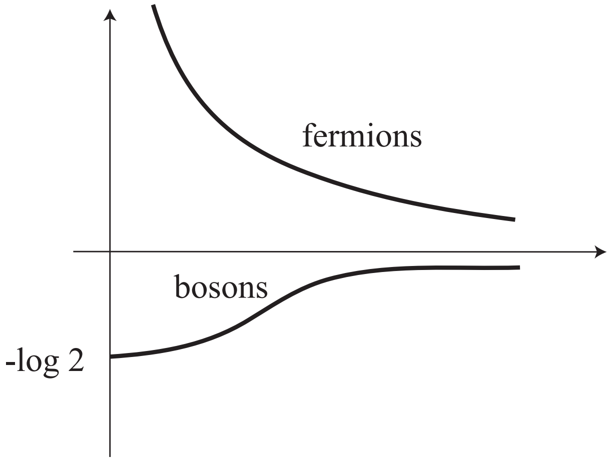
$$\frac{\int d\boldsymbol{p}\, e^{-\beta(p^2/2m)+i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}}{\int d\boldsymbol{p}\, e^{-\beta(p^2/2m)}} = e^{-mk_BTr^2/2\hbar^2}. \tag{6.3.67}$$

Hence,

$$F = 1 \pm e^{-mk_BT(\boldsymbol{r}_1-\boldsymbol{r}_2)^2/\hbar^2}. \tag{6.3.68}$$

(7) If we introduce the effective potential $\phi$ by $F = e^{-\beta\phi}$, we get

$$\beta\phi(\boldsymbol{r}) = -\log[1 \pm e^{-mk_BT(\boldsymbol{r}_1-\boldsymbol{r}_2)^2/\hbar^2}]. \tag{6.3.69}$$



The sketches of the potential are given in the attached figure.

As expected, the effective interaction is attractive for bosons, and repulsive for fermions.

**3.11** [Elementary low temperature formulas for fermions]
The following questions ask for standard elementary calculations, but you should do them once in your life.
(1) Obtain the chemical potential (the Fermi level) to order $T^2$ around $T = 0$.
(2) Obtain the pressure $P$ to order $T^2$ around $T = 0$.

**Solution**
(1) Let us apply

$$\int_0^{+\infty} d\varepsilon\, \phi(\varepsilon) f(\varepsilon) = \int_0^{\mu(T)} \phi(x)dx + (k_BT)^2\zeta(2)\phi'(\mu) + \cdots \tag{6.3.70}$$

to the following calculation:

$$N = \int_0^{\infty} d\varepsilon\, gD_t(\varepsilon)f(\varepsilon) = \int_0^{\mu(T)} dx\, gD_t(x) + \zeta(2) \left.\frac{dgD_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu} (k_BT)^2 + \cdots . \tag{6.3.71}$$

We know the $T = 0$ result:

$$N = \int_0^{\mu(0)} d\varepsilon\, gD_t(\varepsilon). \tag{6.3.72}$$

This determines the Fermi energy $\mu(0)$.

$$\int_0^{\mu(0)} d\varepsilon\, gD_t(\varepsilon) = \int_0^{\infty} d\varepsilon\, gD_t(\varepsilon) f(\varepsilon) = \int_0^{\mu(T)} dx\, gD_t(x) + \zeta(2)\left.\frac{dgD_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu(T)} (k_BT)^2 + \cdots, \tag{6.3.73}$$

so we can expect $\mu(T) = \mu(0) + aT^2 + o[T^2]$:

$$\int_0^{\mu(T)} dx\, gD_t(x) = \int_0^{\mu(0)+aT^2} dx\, gD_t(x) = \int_0^{\mu(0)} dx\, gD_t(x) + aT^2\, gD_t(\mu(0)). \tag{6.3.74}$$

Now, combining (6.3.73) and (6.3.74), we obtain

$$\int_0^{\mu(0)} d\varepsilon\, gD_t(\varepsilon) = \int_0^{\mu(0)} dx\, gD_t(x) + aT^2\, gD_t(\mu(0)) + \zeta(2)\left.\frac{dgD_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu(0)} (k_BT)^2 + \cdots, \tag{6.3.75}$$

or

$$aT^2\, gD_t(\mu(0)) + \zeta(2)\left.\frac{dgD_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu(0)} (k_BT)^2 = 0. \tag{6.3.76}$$

Therefore,

$$a = -\zeta(2)\left.\frac{d\log D_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu(0)} k_B^2 = -\frac{\pi^2}{6}\left.\frac{d\log D_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu(0)} k_B^2. \tag{6.3.77}$$

Thus, the final result is

$$\mu(T) = \mu(0) - \frac{\pi^2}{6}\left.\frac{d}{d\varepsilon}\log D_t(\varepsilon)\right|_{\varepsilon=\mu(0)} (k_BT)^2 + \cdots. \tag{6.3.78}$$

(2) $P$ is required, but it is easier to compute $E$. Utilizing (6.3.70), we get

$$\begin{aligned} E(T) &= \int d\varepsilon\, gD_t(\varepsilon)\varepsilon f(\varepsilon), & (6.3.79)\\ &= \int_0^{\mu} d\varepsilon gD_t(\varepsilon)\varepsilon + \zeta(2)(k_BT)^2\left.\frac{dgD_t(\varepsilon)\varepsilon}{d\varepsilon}\right|_{\mu} + \cdots. & (6.3.80)\end{aligned}$$

We must expand $\mu(T)$ using the result of (1):

$$\begin{aligned} \int_0^{\mu(T)} d\varepsilon gD_t(\varepsilon)\varepsilon &= \int_0^{\mu(0)} d\varepsilon gD_t(\varepsilon)\varepsilon - \frac{\pi^2}{6}\left.\frac{d\log D_t(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\mu(0)} gD_t(\mu(0))\mu(0)(k_BT)^2, & (6.3.81)\\ &= E(0) - \frac{\pi^2}{6} gD_t'(\mu(0))\mu(0)(k_BT)^2 & (6.3.82)\end{aligned}$$

Comparing this with (6.3.80), we can write

$$E(T) = E(0) - \frac{\pi^2}{6} g D_t'(\mu(0))\mu(0)(k_BT)^2 + \zeta(2)(k_BT)^2 \left.\frac{dgD_t(\varepsilon)\varepsilon}{d\varepsilon}\right|_\mu + \cdots, \tag{6.3.83}$$

$$= E(0) + \frac{\pi^2}{6} g D_t(\mu(0))(k_BT)^2. \tag{6.3.84}$$

Since $P = 2E/3V$

$$P(T) = P(0) + \frac{\pi^2}{9V} g D_t(\mu(0))(k_BT)^2. \tag{6.3.85}$$

**3.12** [Derivation of Maxwell's distribution]
Maxwell derived in his *Illustrations of the Dynamical Theory of Gases* (1860) the density distribution function $f(\boldsymbol{v})$ of the gas particle velocity.

Maxwell assumed that orthogonal components of the velocity are statistically independent. This implies that we may write

$$f(\boldsymbol{v}) = \phi_x(v_x)\phi_y(v_y)\phi_z(v_z), \tag{6.3.86}$$

where $\phi_x$, etc., are density distribution function for each component. Maxwell also assumed isotropy. Hence, $f$ is a function of $v^2 \equiv |\boldsymbol{v}|^2$, so we may regard $f(\boldsymbol{v}) \equiv F(v^2)$, and $\phi_x$'s do not depend on suffixes. Let us introduce $\psi(s^2) \equiv \phi_x(s)$. Then, the above functional equation reads

$$F(x + y + z) = \psi(x)\psi(y)\psi(z). \tag{6.3.87}$$

If $F$and $\psi$ are both once differentiable, we obtain

$$F'(x + y + z) = \psi(x)\psi(y)\psi'(z). \tag{6.3.88}$$

Setting $y = z = 0$, we have

$$F(x) = \psi(x)\psi(0)\psi(0), \;\; F'(x) = \psi(x)\psi(0)\psi'(0), \tag{6.3.89}$$

so $F'(x)/F(x)$ must be a constant. This implies that the distribution is Gaussian.
(1) Is there any other solution? If we do not assume the differentiability of $F$ (that is, if we assume only the continuity of $F$), what do you get?[23]
(2) Since we know the result of equilibrium statistical mechanics, if the particle

---

[23] If we do not assume the continuity of $F$, there would be uncountably many solutions.

energy is $E$, then the distribution function is proportional to $e^{-\beta E}$. This is derived from the consistency of mechanics and thermodynamics. On the other hand, the above derivation of the Maxwell distribution uses only the statistical independence of the orthogonal components and its isotropy, and mechanics is never used. Then, this seems to imply that Maxwell's logic determines the form of the kinetic energy $K$ in terms of velocity from statistically natural assumption + thermodynamics; at least $K \propto \boldsymbol{v}^2$ is concluded. This sounds incredible, even if thermodynamics is great. What is wrong? [Hint: think of relativistic case.]

[Comment] Maxwell himself did not like the above derivation we criticize here,[24] so he rederived the distribution a few years later. He this time used the detailed balance argument (as explained in the text). Pay due respect to Maxwell's sound instinct.

**Solution**

(1) Needless to say, if we assume differentiability there is no other solution. Maxwell was correct.

Let us try to solve the problem assuming only continuity (and isotropy). Let us introduce $g = \log F$ and $\phi = \log \psi$; this is admissible because $F$ and $\psi$ are positive. We have

$$g(x+y) = \phi(x) + \phi(y) + \phi(0) = \phi(x+y) + 2\phi(0). \tag{6.3.90}$$

Therefore, if we define $f(x) = \phi(x) - \phi(0)$, then we get

$$f(x) + f(y) = f(x+y). \tag{6.3.91}$$

Since we assume $f$ to be continuous, the solution is $f(x) = cx$ for some constant $c$. Thus, we can get only a Gaussian form.

(2) If we consider the relativistic case, the velocity distribution function reads

$$P(\boldsymbol{v}) \propto \exp(mc^2/\sqrt{1 - v^2/c^2}). \tag{6.3.92}$$

Obviously, it does not have the structure (6.3.86). That is, orthogonal components are not statistically independent, although isotropy is still correct.

**3.13** [2-dimensional neutron system]

$10^{18}$ neutrons are confined in a square of edge length 1 m. If we regard this as a 2-dimensional system, estimate the needed temperature required for this system to be regarded as a classical system.[25]

**Solution**

---

[24] However, even strict Pauli uncritically repeat the above argument in W. Pauli, *Thermodynamics and the Kinetic Theory of Gases* (edited by C. P. Enz), Section 25.

[25] cf ProblWS p176.

Non-classical behavior is observed if the interparticle distance is comparable to the thermal wave length of the particle. This happens when the number density becomes comparable to the so-called quantum number density $n_Q$ (the quantum density must be computed in 2-space):

$$n_Q = 2\pi m k_B T/h^2. \tag{6.3.93}$$

Since $m = 1.65 \times 10^{-27}$kg, $k_B = 1.38 \times 10^{-23}$J/K, $h = 6.63 \times 10^{-34}$ Js, $n/n_Q \sim 1$ implies $T \sim 3.1$K. That is, if the temperature is as high as 30 K, the system behaves classically.

**3.14** [2-dimensional fermion system]
The density of translational states of a 2D fermion system confined in a volume (area) $V$ may be written as $c\, D(\varepsilon) = cV$ with a positive constant $c$.
(1) Find the chemical potential $\mu$ in terms of the number density $\rho$ and (inverse) temperature $\beta$.
(2) In the high density limit, we have $\mu \propto \rho$. Explain why this form is plausible.
(3) What is the classical limit? Does the obtained result consistent with the classical ideal gas result?

**Solution**
(1) We can write

$$\rho = c \int_0^\infty d\varepsilon \frac{1}{e^{\beta(\varepsilon-\mu)} + 1}. \tag{6.3.94}$$

To integrate this, recall:

$$\frac{1}{e^{\beta(\varepsilon-\mu)} + 1} = \frac{d}{d(\beta\mu)} \log[1 + e^{-\beta(\varepsilon-\mu)}]. \tag{6.3.95}$$

Therefore,

$$\rho = c \int_0^\infty d\varepsilon\, \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} = -c \int_0^\infty d\varepsilon \frac{d}{d(\beta\varepsilon)} \log[1 + e^{-\beta(\varepsilon-\mu)}] = (c/\beta) \log[1 + e^{\beta\mu}]. \tag{6.3.96}$$

That is,

$$\mu = \beta^{-1} \log(e^{\beta\rho/c} - 1). \tag{6.3.97}$$

In the high density limit, $\mu \simeq \rho/c$.
(2) The high density limit implies high degeneracy for a Fermi gas.[26]D We may

---

[26] The slope of the cliff of the Fermi distribution in the present case is $4/k_B T$, so you might think it is not sharp. However, the width of the cliff ($\sim 5kBT$) must be compared with the width of the plateau $\mu(0) = k_B T_F$, which becomes indefinitely large as the density increases, so the distribution is after all close to a step function of the low temperature limit.

approximate the distribution as a step function going from 1 to 0 around the Fermi level $\mu$. Therefore, the integration range of (6.3.94) is essentially from 0 to $\mu$. Hence, $\rho = c\mu$ is very plausible.
(3) We certainly have

$$\mu \to k_B T \log(\beta\rho/c); \tag{6.3.98}$$

consistent!

**3.15** [Quantum gas with internal degrees of freedom]
Let us consider a quantum gas consisting of $N$ particles. Individual particles have internal states consisting of two levels: the ground state and the non-degenerate excited state with energy $\varepsilon$ ($> 0$).
(1) Suppose the particles are fermions. How does the Fermi energy $\mu_F$ (i.e., the chemical potential) behave as a function of $\varepsilon$?
(2) Suppose the particles are bosons. How does the Bose-Einstein critical temperature $T_c$ depends on $\varepsilon$? Give a clear argument even if it is qualitative.
**Solution**
(1) The Fermi energy $\mu$ is determined by

$$N = \int dE\, \mathcal{D}(E) \frac{1}{e^{\beta(E-\mu)} + 1}. \tag{6.3.99}$$

If $\varepsilon$ is increase, then the occupation number of the one-particle states with internal excitation. If the number of particles is constant, then to accommodate these internally non-excited particles, the Fermi level must be increased. Therefore, $\mu$ is an increasing function of $\varepsilon$.
(2) Consider the total number of internally excited particles (note that $\mu = 0$):

$$N_1 = \int dE\, \mathcal{D}(E) \frac{1}{e^{\beta E} - 1}. \tag{6.3.100}$$

If $\varepsilon$ is increased, $N_1$ decreases, so this favors the formation of condensate. That is, $T_c$ increases with $\varepsilon$.

**3.16** [Zeemann splitting]
The outer shell of an ion has a magnetic moment $\mu_B$ of 1 Bohr magneton. In a magnetic field $B$ this outer shell state splits into two energy states with energies $E = E_0 \pm \mu_B B$. Let $n_u$ (resp., $n_d$) be the occupancy number of up-spin (resp., down-spin) states. Then the magnetization reads $M = \mu_B(n_u - n_d)$. You may ignore the electron-electron interactions.
(1) Find $\langle M \rangle$ and $\langle N \rangle$ ($N = n_u + n_d$) with the aid of the grand canonical formalism.

(2) Find the magnetization when the outer shell has one electron for each ion. Compare the result with the result of (1) for $\mu = E_0$.[27]

**Solution**

(1) Since

$$n_u = \frac{1}{e^{\beta(E_0+\mu_B B-\mu)}+1},\ n_d = \frac{1}{e^{\beta(E_0-\mu_B B-\mu)}+1}, \tag{6.3.101}$$

$\langle M\rangle$ and $\langle N\rangle$ can be written down immediately.

(2) Define $x = e^{\beta(\varepsilon-\mu)}$ and $y = e^{\beta\mu_B H}$. Then, $\langle N\rangle = 1$ can be written as

$$1 = n_u + n_d = \frac{2 + x(y+1/y)}{1 + x(y+1/y) + x^2}. \tag{6.3.102}$$

If we set $x = 1$, this equality holds for any $y$. Therefore, $E_0 = \mu$ is the condition, and

$$\langle M\rangle = \mu_B\left(\frac{1}{1+xy} - \frac{1}{1+x/y}\right) = \mu_B\frac{1-y^2}{1+y^2}. \tag{6.3.103}$$

**3.17** [Electron paramagnetism]

Due to the spin, each electron in a magnetic field $B$ (assumed to be pointing the $z$-direction) has the potential energy $\pm\tilde{\mu}B$. Let $D_t(\varepsilon)$be the one-particle translational density of states (however, the electrons may be in a crystal field, so we do not specify its form).

(1) The magnetization $M$ of this system $M$ is the expectation of the magnetic moment due to electron spins. Express $M$ in terms of $D_t(\varepsilon \pm \tilde{\mu}B)$.

(2) Express the magnetic susceptibility $\chi$ in terms of $D_t'(\varepsilon)$, assuming that $\tilde{\mu}B$ is not too large.

(3) Obtain $\chi$ to order $T^2$ around $T = 0$ with the aid of $\log D_t(\varepsilon)$.

**Solution**

(1) The contribution of the up-spin electrons to the magnetization is

$$M_+ = \tilde{\mu}\int d\varepsilon\, D_t(\varepsilon)\frac{1}{e^{\beta(\varepsilon-\tilde{\mu}B-\mu)}+1} = \tilde{\mu}\int d\varepsilon\, D_t(\varepsilon+\tilde{\mu}B)\frac{1}{e^{\beta(\varepsilon-\mu)}+1}. \tag{6.3.104}$$

We can easily obtain the analogous formula for down-spin electron, so we get

$$M = M_+ + M_- = \tilde{\mu}\int d\varepsilon\,[D_t(\varepsilon+\tilde{\mu}B) - D_t(\varepsilon-\tilde{\mu}B)]\frac{1}{e^{\beta(\varepsilon-\mu)}+1}. \tag{6.3.105}$$

---

[27] UIUC QualFall 95

(2) $\tilde{\mu}B$ is assumed to be not too large, we may Taylor expand as

$$M = \tilde{\mu} \int d\varepsilon \, [D_t(\varepsilon + \tilde{\mu}B) - D_t(\varepsilon - \tilde{\mu}B)] \frac{1}{e^{\beta(\varepsilon-\mu)}+1} = 2\tilde{\mu}^2 B \int d\varepsilon \, D_t'(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)}+1}. \tag{6.3.106}$$

(3) From the definition

$$\chi = \left.\frac{\partial M}{\partial B}\right|_T = 2\tilde{\mu}^2 \int_0^\infty d\varepsilon \, D_t'(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)}+1}. \tag{6.3.107}$$

To expand this around $T = 0$ we use, setting $\phi = 2\tilde{\mu}^2 D_t'(\varepsilon)$,

$$\int_0^\infty d\varepsilon \, \phi(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)}+1} = \int_0^\mu d\varepsilon \, \phi(\varepsilon) + (k_BT)^2 \zeta(2) \phi'(\mu). \tag{6.3.108}$$

That is,

$$\chi = 2\tilde{\mu}^2 D_t(\mu) + 2\tilde{\mu}^2 (k_BT)^2 \zeta(2) D_t''(\mu). \tag{6.3.109}$$

This is, however, not yet the final result. Since

$$\mu = \mu_0 - \zeta(2) \left.\frac{d}{d\varepsilon} \log D_t(\varepsilon)\right|_{\varepsilon=\mu_0}, \tag{6.3.110}$$

(6.3.109) reads

$$\chi = 2\tilde{\mu}^2 D_t(\mu_0) - 2\tilde{\mu}^2 (k_BT)^2 \zeta(2) D_t'(\mu_0) \left.\frac{d}{d\varepsilon} \log D_t(\varepsilon)\right|_{\varepsilon=\mu_0} + 2\tilde{\mu}^2 (k_BT)^2 \zeta(2) D_t''(\mu). \tag{6.3.111}$$

This can be streamlined to the following form:

$$\begin{aligned} \chi &= 2\tilde{\mu}^2 D_t(\mu_0) + 2\tilde{\mu}^2 (k_BT)^2 \zeta(2) D_t(\mu_0) \left.\frac{d^2}{d\varepsilon^2} \log D_t(\varepsilon)\right|_{\varepsilon=\mu_0} \\ &= 2\tilde{\mu}^2 D_t(\mu_0) \left[ 1 + \zeta(2)(k_BT)^2 \left.\frac{d^2}{d\varepsilon^2} \log D_t(\varepsilon)\right|_{\varepsilon=\mu_0} \right]. \end{aligned} \tag{6.3.112}$$

**3.18** [Do we have only to treat the ground state special below $T_c$?]

For a bose gas in 3-space we know the following integral expression is not always correct:

$$\langle N\rangle = \sum_{i=0}^{\infty}\langle \hat{n}_i\rangle = \int_0^{\infty} d\varepsilon\, D_t(\varepsilon). \tag{6.3.113}$$

It is because the expression ignores a large number of particles in the one-particle ground state. Thus, we are taught that if we count the number $N_0$ of the particles occupying the one-particle ground state and if we add this to $N_1$, then the number of particles in the system may be expressed correctly. However, ther may be the people who are not so convinced yet: why only ground state? Don't we have to consider the first excited state? Don't we actually have to perform the following calculation $\cdots$:

$$\frac{\langle N\rangle}{V} = \frac{1}{V}\langle \hat{n}_0\rangle + \frac{1}{V}\langle \hat{n}_1\rangle + \frac{1}{V}\int_0^{\infty} d\varepsilon\, D_t(\varepsilon). \tag{6.3.114}$$

Let us perform a slightly more honest calculation (to recognize clearly that Einstein is always correct!):

(1) Our energy coordinate convention is that the ground state is always 0: $\varepsilon_0 = 0$. Let us assume that the system is a cube of edge length $L$: $V = L^3$. The lowest excited one-particle state energy $\varepsilon_1$ as a function of $V$.

(2) Compare the occupation number of the one-particle ground state and the one particle first excited states (which is triply degenerate). That is, compute the ratio $(\langle \hat{n}_0\rangle/(\langle \hat{n}_1\rangle + \langle \hat{n}_2\rangle + \langle \hat{n}_3\rangle) = \langle \hat{n}_0\rangle/3\langle \hat{n}_1\rangle$ for a very small negative chemical potential $\mu$[28] required by the Bose-Einstein condensation. How big is it as a function of $V$?

(3) We just saw in (2) except for $\langle \hat{n}_0\rangle$ other expectation values are not extensive. That is, the ground state is really special. Excited states cannot contribute an extensive quantity unless infinitely many of them are collected. Explain that the contribution of all the excited states may be obtained accurately by replacing the summation with integration (as usual).

**Solution**

(1) This calculation is just as we did in Chapter 1:

$$\varepsilon_{\boldsymbol{n}} = \frac{h^2}{8mV^{2/3}}(n_x^2 + n_y^2 + n_z^2), \tag{6.3.115}$$

where $n$'s are positive integer quantum numbers. Therefore, the energy difference between the ground state and the first excited state is

$$\Delta\varepsilon = 3\frac{h^2}{8mV^{2/3}}. \tag{6.3.116}$$

---

[28] which is not zero, because the system is finite.

This is $\varepsilon_1$ according to our convention.
(2) We have

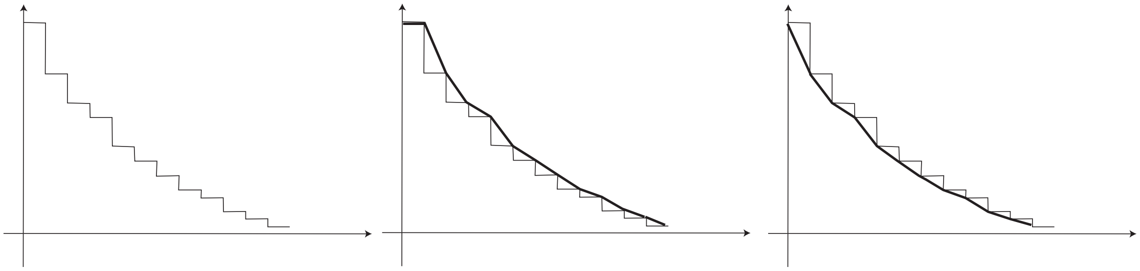
$$\frac{\langle \hat{n}_0 \rangle}{3\langle \hat{n}_1 \rangle} = \frac{e^{\beta(\varepsilon_1 - \mu)} - 1}{e^{-\beta\mu} - 1}. \tag{6.3.117}$$

We know below $T_c$ $\beta\mu = O[N^{-1}]$ ($< 0$). Furthermore, we know $\varepsilon_1 = O[V^{-2/3}]$. Since $T > 0$ is a fixed temperature, however small it is (or however large $\beta$ is), if we take a sufficiently large $V$, we may regard $\beta\varepsilon_1$ to be sufficiently small ($\beta\mu$ is much smaller than this), so we may expand as

$$\frac{\langle \hat{n}_0 \rangle}{3\langle \hat{n}_1 \rangle} = \frac{\varepsilon_1 - \mu}{-3\mu} = \frac{1}{3}(1 - \varepsilon_1/\mu) = O[N^{1/3}] \gg 1. \tag{6.3.118}$$

Thus, we see that only the one-particle ground state is occupied by an extensive number of particles; any finite some of the occupation numbers of one-particle excited states is far less than $N_0$ for large systems.
(3) Let $\{f(i)\}$ be a monotone decreasing sequence of positive integers and assume $\sum f(i)$ converge. Define monotone decreasing (piecewise linear ) functions $f_L(x)$ as $f_L(i-1) = f(i)$ for $i = 1, 2, \cdots$ and $f_U(x)$ as $f_U(i) = f(i)$ for $i = 1, 2, \cdots$ and $f_U(0) = f_U(1)$ (see the graphs below). Then,

$$\int_0^\infty f_L(x)dx \le \sum_{i=1}^\infty f(\varepsilon) \le \int_0^\infty f_U(x)dx. \tag{6.3.119}$$



The thick curve in the center is $f_U$; that in the right is $f_L$.

As can easily be seen from the figure

$$\int_0^\infty f_U(x)dx - \int_0^\infty f_L(x)dx < f(\varepsilon_1). \tag{6.3.120}$$

Therefore, the difference divided by $V$ is extremely small.

**3.19** [Ideal boson gas slightly warmer than $T_c$]
Fig. 3.4.2 illustrates that the specific heat $C_v$ of the ideal Bose gas has a cusp. Let us demonstrate this. To compute $C_v$ we need the internal energy of the system. Let us compute it.
(1) What is the internal energy below $T_c$? (This is an easy question.)
(2) If we compute the internal energy assuming $\mu = 0$ and write its value as $E_0$, show

$$\frac{\partial E_0}{\partial \mu} \simeq \frac{3}{2} N_0(T), \tag{6.3.121}$$

where

$$N_0(T) \equiv N(T,0) = \frac{V}{h^3}\int_0^\infty \frac{1}{e^{\beta p^2/2m}-1} 4\pi p^2 dp. \tag{6.3.122}$$

Therefore, for $T$ ($> T_c$) we could approximate the true internal energy at $T$ as $E(T) = E_0 + (3/2)N_0(T)\mu$. This implies that to obtain $E$ as a function of $T$, we need $\mu$ as a function of $T$. To this end let us write the number of particles for $T > T_c$ ($\mu < 0$) as

$$N = N_0(T) + \frac{V}{h^3}\int_0^\infty \left\{ \frac{1}{e^{\beta(p^2/2m-\mu)}-1} - \frac{1}{e^{\beta p^2/2m}-1} \right\} 4\pi p^2 dp. \tag{6.3.123}$$

(3) Show that we may approximate the second term of (6.3.123) as

$$\int_0^\infty \left\{ \frac{1}{e^{\beta(\varepsilon-\mu)}-1} - \frac{1}{e^{\beta\varepsilon}-1} \right\} \sqrt{\varepsilon} d\varepsilon \simeq k_BT\mu \int_0^\infty d\varepsilon \frac{1}{\sqrt{\varepsilon}(\varepsilon+|\mu|)} = -\pi k_BT\sqrt{|\mu|}. \tag{6.3.124}$$

Do not forget that $\mu < 0$. [This allows us to obtain $\mu$ in terms of $N_0(T)$ which is obtainable from (6.3.122) as a function of $T$.]

**Solution**
(1) Simply copy the formula:

$$E = \int d\varepsilon g D_t(\varepsilon)\frac{\varepsilon}{e^{\beta\varepsilon}-1} = \frac{3}{2}k_BTV\left(\frac{2\pi m k_BT}{h^2}\right)^{3/2}\zeta(5/2) \propto VT^{5/2}, \tag{6.3.125}$$

(2) Since

$$E = \frac{Vm^{3/2}}{2^{1/2}\pi^2\hbar^3}\int_0^\infty \frac{\varepsilon^{3/2}}{e^{\beta(\varepsilon-\mu)}-1} d\varepsilon, \tag{6.3.126}$$

we have

$$\frac{\partial E}{\partial \mu} = \frac{Vm^{3/2}}{2^{1/2}\pi^2\hbar^3}\int_0^\infty \varepsilon^{3/2}\frac{\partial}{\partial\mu}\frac{1}{e^{\beta(\varepsilon-\mu)}-1} d\varepsilon, \tag{6.3.127}$$

$$= -\frac{Vm^{3/2}}{2^{1/2}\pi^2\hbar^3}\int_0^\infty \varepsilon^{3/2}\frac{\partial}{\partial\varepsilon}\frac{1}{e^{\beta(\varepsilon-\mu)}-1}d\varepsilon, \tag{6.3.128}$$

$$= \frac{Vm^{3/2}}{2^{1/2}\pi^2\hbar^3}\int_0^\infty \left(\frac{\partial}{\partial\varepsilon}\varepsilon^{3/2}\right)\frac{1}{e^{\beta(\varepsilon-\mu)}-1}d\varepsilon, \tag{6.3.129}$$

$$= \frac{3}{2}\frac{Vm^{3/2}}{2^{1/2}\pi^2\hbar^3}\int_0^\infty \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-\mu)}-1}d\varepsilon. \tag{6.3.130}$$

Comparing this with (6.3.122), we get (6.3.121).
(3) The first approximate relation in (6.3.124) is due to simple Taylor expansion. The integral in the second formula could be computed with the aid of complex analysis (you must respect the branching due to $\sqrt{\varepsilon}$), but an easier way may be to set $\sqrt{\varepsilon} = x$:

$$\int_0^\infty d\varepsilon\frac{1}{\sqrt{\varepsilon}(\varepsilon+|\mu|)} = 2\int_0^\infty dx\frac{1}{x^2+|\mu|} = \int_{-\infty}^\infty \frac{dx}{x^2+|\mu|}. \tag{6.3.131}$$

**3.20** [Bose-Einstein condensation in a harmonic trap]
Let us consider an ideal bose gas consisting of $N$ particles confined in a 3D harmonic potential.[29] It is hard to treat this in terms of the canonical ensemble, so we discuss this with the aid of the grand canonical theory; if $N$ is larger than $10^3$, then $\log N/N$[30] is not large, so this approach must not be bad.
(1) Suppose the angular frequency of the trapped boson is $\omega_t$. Find the density $D(\varepsilon)$ of one-particle state as a function of energy $\varepsilon$. Measure the energy from the ground state and ignore the zero-point energy.
(2) Find the number of particles $N_1$ in the non-condensate as a function of the chemical potential. Show that the integral (or $N_1$) is bounded from above in 3-space (no explicit integration required). Thus, Bose-Einstein condensation is possible in this system.
(3) The number of particles occupying the one-particle ground state approaches zero as

$$N_0(T) = N\left(1-\left(\frac{T}{T_c}\right)^\gamma\right), \tag{6.3.132}$$

when $T \nearrow T_c$. Find $\gamma$.
(4) Find $T_c$ as a function of $N$. For $N = 3000$, and $\omega_t = 10^3$ rad/s estimate $T_c$. (Use $\zeta(3) = 1.2020569031595\cdots$.)

---

[29] Actually, BEC is observed in a collection of Rb atoms confined in a (not spherically symmetric) 3D harmonic potential.

[30] It is emphasized again that the error is not of order $\sqrt{N}/N = 1/\sqrt{N}$, but $\log N/N$.

(5) If a 2D harmonic potential can trap 2D Bose gas,[31] can we observe Bose-Einstein condensation?

**Solution**

(1) The one particle energy in a 3D harmonic potential may be written as $\varepsilon = \hbar\omega_t(n_x + n_y + n_z)$, where $n_x$ are nonnegative integers. The number of energy levels up to energy $\varepsilon$ is the volume of a cone $x + y + z \in [0, \varepsilon/\hbar\omega_t]$ within the first octant, we get

$$\int_0^\varepsilon D(\varepsilon')d\varepsilon' = \frac{1}{3!}\left(\frac{\varepsilon}{\hbar\omega_t}\right)^3. \tag{6.3.133}$$

That is,

$$D(\varepsilon) = \frac{\varepsilon^2}{2(\hbar\omega_t)^3}. \tag{6.3.134}$$

(2) By using $D$ obtained in (1), the number of particles in the excited states is given by

$$N_1 = \int_0^\infty d\varepsilon D(\varepsilon)\frac{1}{e^{\beta(\varepsilon-\mu)} - 1} = \int_0^\infty d\varepsilon' \frac{\varepsilon'^2}{2(\hbar\omega_t)^3}\frac{1}{e^{\beta(\varepsilon'-\mu)} - 1}. \tag{6.3.135}$$

This is an increasing function of $\mu$, so by setting $\mu = 0$ an upper bound of $N_1$ may be obtained:

$$N_1 \le \frac{1}{2}\left(\frac{k_BT}{\hbar\omega_t}\right)^3 \int_0^\infty dx \frac{x^2}{e^x - 1}. \tag{6.3.136}$$

This integral from 1 to $\infty$ converges and

$$\int_0^1 dx \frac{x^2}{e^x - 1} \le \int_0^1 x dx = \frac{1}{2}. \tag{6.3.137}$$

Therefore, $N_1$ is bounded from above. We may estimate it as

$$N_1(T) \le AT^3, \tag{6.3.138}$$

where $A$ ($> 0$) is an appropriate constant. Hence, BEC must occur.

(3) If $\mu = 0$, we know from (2) that $N_1(T) \propto T^3$. Therefore, $\gamma = 3$.

(4) To estimate $T_c$ we need the value of $A$ in $N_1 = AT_c^3$: at $T_c$

$$N = N_1 = \int_0^\infty d\varepsilon' \frac{\varepsilon'^2}{2(\hbar\omega_t)^3}\frac{1}{e^{\beta\varepsilon'} - 1} = \frac{1}{2}\left(\frac{k_BT}{\hbar\omega_t}\right)^3 \int_0^\infty \frac{x^2}{e^x - 1}dx = \zeta(3)\left(\frac{k_BT}{\hbar\omega_t}\right)^3. \tag{6.3.139}$$

---

[31] This is virtually realized on graphene.

Therefore, we conclude

$$T_c = (1.2)^{-1/3}\frac{\hbar\omega_t}{k_B}N^{1/3} = 9.02 \times 10^{-8}\ \mathrm{K}. \tag{6.3.140}$$

(5) For 2D harmonic potential

$$D(\varepsilon) = \varepsilon/(\hbar\omega_t)^2. \tag{6.3.141}$$

Therefore, $N_1$ is computed as

$$N_1 = \int_0^\infty d\varepsilon' \frac{\varepsilon'}{(\hbar\omega_t)^2}\frac{1}{e^{\beta(\varepsilon'-\mu)}-1} \le \int_0^\infty d\varepsilon' \frac{\varepsilon'}{(\hbar\omega_t)^2}\frac{1}{e^{\beta\varepsilon'}-1}, \tag{6.3.142}$$

which is bounded from above (bounded by some $AT^2$). Therefore, BEC can occur in 2D if harmonically bound (quite different from the free space).

**3.21** [Expanding universe]
At present, the cosmic background radiation is at 3 K. Suppose the volume of the universe doubles adiabatically. What is the temperature of the cosmic background radiation after this expansion?

**Solution**
We know the entropy of the radiation field is $S \propto VT^3$. If the system expand quasistatically, the entropy is constant, so $(2V)T^3 = V3^3$. That is, $T = 3/2^{1/3}$ K. Actually, the process may not be quasistatic, so this estimate must be the lower bound.[32]

**3.22** [Specific heat of hydrogens]
Consider a 1 mole of ideal gas at 10 K consisting of pure HD, pure HT or pure DT. Whose specific heat $C_V$ is the largest? Give your answer without detailed computation. You may assume that the length of the chemical bonds are all the same.

**Solution**
We may totally ignore the the contribution of oscillations. There is no difference in the contribution of translational motions. These are all heteronuclear molecules, so we need not worry about spin-rotation coupling. Therefore, we have only to pay attention to the rotational contributions. The molecules with the largest moment of inertia is the easiest to excite, so their rotational specific heat is the largest (notice that the peak of the rotational specific heat occurs around 40 K or above). Therefore,

---

[32] This is virtually a monatomic ideal gas problem, but do not forget that the particles are superrelativistic.

the specific heat of DT must be the largest around 10K. This is indeed the case.

**3.23** [Internal degree of freedom of heavy hydrogen]
The potential energy function describing the chemical bond in a heavy hydrogen $D_2$ may be approximately described as

$$\phi(r) = \varepsilon \left[ e^{-2(r-d)/a} - 2e^{-(r-d)/a} \right], \tag{6.3.143}$$

where $\varepsilon = 7 \times 10^{-19}$ J, $d = 8 \times 10^{-11}$ m and $a = 5 \times 10^{-11}$ m.
(1) Evaluate the smallest energy required to excite the rotational motion, and estimate the temperature $T_r$ for which the rotation starts to contribute significantly.
(2) Evaluate the smallest energy required to excite the vibrational motion, and estimate the temperature $T_rv$ for which the vibration starts to contribute significantly.

**Solution**
(1) The moment of inertia is

$$I = \frac{1}{2}md^2 = (1/2) \times (1.66 \times 10^{-27}) \times (8 \times 10^{-11})^2, \tag{6.3.144}$$

so $\Theta_r = \hbar^2/2k_BI \simeq 150$K.
(2) The vibrational quantum is $\hbar\omega/k_B \simeq 6300$K.

**3.24** [Computation of inertial moment tensor]
Obtain the moment of inertia tensor for $CH_3D$ around its center of mass, and compute its rotational partition function classically. [You have only to state your strategy without actually estimating the components of the tensor.]

**Solution**
The inertial moment tensor around the center of mass $A_{CM}$ and that $A_O$ around an arbitrary point $O$ is related as

$$A = A_O - A_{CM}, \tag{6.3.145}$$

where

$$A_{CM} = M\left[ r_{CM}^2 I - \boldsymbol{r}_{CM}\boldsymbol{r}_{CM}^T \right]. \tag{6.3.146}$$

Here, $M$ is the total mass, $\boldsymbol{r}_{CM}$ is the CM coordinate vector relative to $O$. $A_O$ is computed as

$$A_O = \sum_i m_i \left[ r_i^2 I - \boldsymbol{r}_i\boldsymbol{r}_i^T \right], \tag{6.3.147}$$

where $m_i$ is the mass of atom $i$, $\boldsymbol{r}_i$ it its position vector relative to $O$.

Let us compute the inertial moment tensor of $CH_3D$. The molecule may be considered to be a regular tetrahedron, so it is convenient to compute the inertial moment tensor around the central carbon. H and D sits at the apices of the tetrahedron and the cC is at its geometrical center. Take the position of the carbon as $O$. Let us write the length of the CH (or CD) bond as $r_0$. Let us take $r_0/\sqrt{3}$ as the length unit. Three H's are placed at $(1,-1,-1)$, $(-1,-1,1)$ and $(-1,1,-1)$, and D at $(1,1,1)$. Then, we have

$$A_O = \begin{pmatrix} 10 & -1 & -1 \\ -1 & 10 & -1 \\ -1 & -1 & 10 \end{pmatrix} m_H^2(r_0^2/3). \tag{6.3.148}$$

The position of the center of mass is $\boldsymbol{r}_{CM} = (1,1,1)/17$, so

$$A_{CM} = \frac{1}{17}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} m_H^2(r_0^2/3). \tag{6.3.149}$$

Combining these results, we have

$$A = \begin{pmatrix} 10-2/17 & -16/17 & -16/17 \\ -16/17 & 10-2/17 & -16/17 \\ -16/17 & -16/17 & 10-2/17 \end{pmatrix} m_H^2(r_0^2/3). \tag{6.3.150}$$

To compute the rotational partition function, we have only to compute the determinant of this matrix.

## 6.4 Problems for Chapter 4

**4.1**. [Kac potential]
There is an imperfect classical gas with a binary potential $\phi$ given by

$$\phi(|\boldsymbol{r}|) = \begin{cases} \infty & \text{if } |\boldsymbol{r}| < a, \\ -\varepsilon/l^3 & \text{if } a \le |\boldsymbol{r}| < l, \\ 0 & \text{otherwise.} \end{cases} \tag{6.4.1}$$

Here, $\varepsilon > 0$, $a$ is a length scale of atom size, and the $l \to \infty$ limit is taken. (This is an example of the Kac potential.)
(1) Compute the second virial coefficient (in the $l \to \infty$ limit).
(2) Compute the Joule-Thomson coefficient $(\partial T/\partial P)_H$, where $H$ is enthalpy. The reader may assume that the heat capacity $C_P$ under constant pressure is a constant and is known.

**Solution**
Since

$$B(T) = \frac{1}{2}\int_0^\infty \left[1 - e^{-\beta\phi(r)}\right] 4\pi r^2 dr, \tag{6.4.2}$$

we have

$$\begin{aligned} 2B(T) &= \int_0^a 4\pi r^2 dr + \int_a^l (1 - e^{\beta\varepsilon/l^3}) 4\pi r^2 dr, & (6.4.3)\\ &\to \frac{4}{3}\pi a^3 - \frac{4}{3}\pi\beta\varepsilon. & (6.4.4)\end{aligned}$$

That is,

$$B(T) = \frac{2\pi}{3}(a^3 - \beta\varepsilon). \tag{6.4.5}$$

(2)

$$\left.\frac{\partial T}{\partial P}\right|_H = \frac{\partial(T,H)}{\partial(T,P)}\frac{\partial(T,P)}{\partial(P,H)} = -\left.\frac{\partial H}{\partial P}\right|_T \Big/ \left.\frac{\partial H}{\partial T}\right|_P. \tag{6.4.6}$$

$(\partial H/\partial T)_P = C_P$ and

$$\left.\frac{\partial H}{\partial P}\right|_T = T\left.\frac{\partial S}{\partial P}\right|_T + V = -T\left.\frac{\partial V}{\partial T}\right|_P + V. \tag{6.4.7}$$

Since the equation of state is

$$PV = Nk_BT\left(1 + B(T)\frac{N}{V} + \cdots\right), \tag{6.4.8}$$

we obtain

$$\frac{V}{N} = \frac{k_B T}{P} + B + O[\rho], \tag{6.4.9}$$

where $\rho$ is the number density, and

$$\left.\frac{\partial V}{\partial T}\right|_P = \frac{N k_B}{P} + N B' + \cdots. \tag{6.4.10}$$

Therefore, (again higher order terms in $\rho$ are ignored)

$$T\left.\frac{\partial V}{\partial T}\right|_P = V + TNB' + \cdots, \tag{6.4.11}$$

so the Joule-Thomson coefficient reads

$$\left.\frac{\partial T}{\partial P}\right|_H = \frac{N}{C_P}\left(T\frac{dB(T)}{dT} - B\right). \tag{6.4.12}$$

Now, introduce $B$ obtained in (1) and we get

$$\left.\frac{\partial T}{\partial P}\right|_H = \frac{2\pi N}{3C_P}\left(2\frac{\varepsilon}{k_B T} - a^3\right). \tag{6.4.13}$$

From this we see that if the temperature is sufficiently low, we can cool the gas using the Joule-Thomson effect (as the ordinary gases).

**4.2** [van der Waals equation of state]
(1) Show that the critical point is defined by

$$\left.\frac{\partial P}{\partial V}\right|_T = \left.\frac{\partial^2 P}{\partial V^2}\right|_T = 0. \tag{6.4.14}$$

(2) For the van der Waals equation of state, find the universal ratio $P_c V_c / k_B T_c$.
(3) Obtain the reduced equation of state $P_r = f(V_r, T_r)$ for the van der Waals gas. Here, $P_r = P/P_c$, $V_r = V/V_c$ and $T_r = T/T_c$ are reduced variables. [The reader can work with a 1 mole gas.]
(4) Near the critical point $P_r - 1$ may be expanded in powers of $T_r - 1$ and $n_r - 1$, where $n_r = 1/V_r$ is the reduced number density. Find the coefficients $A$ - $C$ (we will see a close relation of this to the Landau theory of phase transition later).

$$P_r - 1 = A(T_r - 1) + B(T_r - 1)(n_r - 1) + C(n_r - 1)^3 + \cdots. \tag{6.4.15}$$

(5) For hydrogen gas $H_2$, $b = 26.61$ cm$^3$/mol. This volume can be interpreted as the smallest volume that the molecules can be squeezed into. Assuming that Avogadro's constant is known, estimate Planck's constant (use dimensional analysis to guess the atom size in terms of electron charge $e$, mass $m$, $h$ and $4\pi\varepsilon_0$).

**Solution**

(1) The condition for the criticality is for the van der Waals loop to become a single point. Therefore, the critical point corresponds to the inflection point of the $PV$-curve (the point where te local max and min points coalesce). This implies the two conditions stated in the problem.

(2) The equation of state we start with is

$$P = \frac{Nk_BT}{V - Nb} - \frac{aN^2}{V^2}, \tag{6.4.16}$$

and the two conditions in (1) read

$$\begin{aligned} -\frac{Nk_BT_c}{(V_c - Nb)^2} + 2\frac{aN^2}{V_c^3} &= 0, && (6.4.17) \\ 2\frac{Nk_BT_c}{(V_c - Nb)^3} - 6\frac{aN^2}{V_c^4} &= 0. && (6.4.18) \end{aligned}$$

Taking the ratio of these two equations, we get $V_c - Nb = 2V_c/3$. That is, $V_c = 3Nb$. From the first equality (6.4.17)

$$\frac{k_BT_c}{(2bN)^2} = 2\frac{aN}{(3bN)^3} \ \Rightarrow\ k_BT_c = \frac{8a}{27b}. \tag{6.4.19}$$

Now, with the aid of the equation of state, we get

$$P_c = \frac{N(8a/27b)}{2Nb} - \frac{aN^2}{9N^2b^2} = \frac{a}{27b^2}. \tag{6.4.20}$$

Combining all the results, we get

$$\frac{P_cV_c}{Nk_BT_c} = \frac{(a/27b^2)(3Nb)}{N(8a/27b)} = \frac{3}{8}. \tag{6.4.21}$$

That is, unless the ratio is 3/8, a gas does not obey the van der Waals equation of state.

(3) Introducing $P = (a/27b^2)P_r$, $k_BT = k_BT_r(8a/27b)$ and $V = (3b)V_r$ into the van der Waals equation of state, we get

$$P_r\frac{a}{27b^2} = \frac{NT_r(8a/27b)}{3bNV_r - Nb} - \frac{aN^2}{9N^2b^2V_r^2}, \tag{6.4.22}$$

that is,

$$P_r = \frac{(8/3)T_r}{V_r - (1/3)} - \frac{3}{V_r^2}. \tag{6.4.23}$$

(4) If you are <u>confident</u> about your analytical muscle, you may leave all the following calculation to, e.g., Mathematica.

Using $P_r = \pi + 1$, $T_r = \tau + 1$, and $V_r = 1/(1+n)$, we can rewrite (6.4.23) as

$$\begin{aligned}
\pi + 1 &= \frac{(8/3)(1+\tau)}{1/(n+1) - 1/3} - 3(1+n)^2 = \frac{8(1+\tau)(1+n)}{2-n} - 3(1+n)^2, && (6.4.24)\\
&= 4(1+\tau)(1+n)\left(1 + \frac{n}{2} + \frac{n^2}{4} + \cdots\right) - 3(1+n)^2, && (6.4.25)\\
&= 4(1+\tau+n+n\tau)\left(1 + \frac{n}{2} + \frac{n^2}{4} + \frac{n^3}{8} + \cdots\right) + (3+6n+3n^2), && (6.4.26)\\
&= 4\left(1 + \tau + \frac{3n}{2} + \frac{3}{2}n\tau + \frac{3}{4}n^2 + \frac{3}{4}n^2\tau + \frac{3}{8}n^3 + \cdots\right) - 3 - 6n - 3n^2, && \\
& && (6.4.27)\\
&= 1 + 4\tau + 6n\tau + \frac{3}{2}n^3 + 3n^2\tau + \cdots. && (6.4.28)
\end{aligned}$$

That is,

$$\pi = 4\tau + 6n\tau + \frac{3}{2}n^3 + \cdots. \tag{6.4.29}$$

This implies that $A = 4$, $B = 6$ and $C = 3/2$.

(5) The radius of the atom may be dimensional-analytically estimated as follows: $[m] = M$, $[e^2/4\pi\varepsilon_0] = ML^3T^{-2}$ and $[h] = ML^2T^{-1}$, so the quantity with the dimension of length that can be constructed from these quantities is

$$[(h/m)^2/(e^2/4\pi m\varepsilon_0)] = (L^2T^{-1})^2/(L^3/T^{-2}) = L. \tag{6.4.30}$$

That is, the radius of an atom may be evaluated as $a = 4\pi\varepsilon_0 h^2/me^2$. This $a$ is something like $(b/N_A)^{1/3}$. Therefore,

$$h = \sqrt{\frac{me^2}{4\pi\varepsilon_0}\left(\frac{b}{N_A}\right)^{1/3}} = 2.2 \times 10^{-34}\ \text{Js}. \tag{6.4.31}$$

This is not very bad (taking the crudeness of the argument into account, it is not at all bad).

**4.3** [The free energy of the van der Waals gas]
The Helmholtz free energy of the van der Waals gas may be expressed as

$$A = -Nk_BT\left\{\log\left[\frac{n_Q}{N}(V-Nb)\right]+1\right\} - \frac{aN^2}{V}. \tag{6.4.32}$$

(1) Comparing this with the free energy formula for the ideal gas, explain why this form is natural.
(2) Compute the internal energy and the entropy of the van der Waals gas.

**Solution**
(1) If we forget about the effect of the attractive forces, the fundamental idea of the van der Waals gas is the ideal gas in the effective volume (= free volume) $V - Nb$, that is, the actual space $-$ the excluded volume due to molecules. The first term of $A$ is, as you can see by comparing it with the ideal gas formula, just the term for such an ideal gas. The second term is the effect of attractive forces: since $\partial A/\partial V = -P$, the effect of the attractive forces may be computed as (integrating $-PdV$)

$$-\int_\infty^V \left[-\frac{aN^2}{V^2}\right] = -\frac{aN^2}{V}. \tag{6.4.33}$$

(2)

$$S = -\frac{A}{T} = N\left\{\log\left[\frac{n_Q}{N}(V-Nb)\right]+\frac{5}{2}\right\}. \tag{6.4.34}$$

$E = A + ST$ is

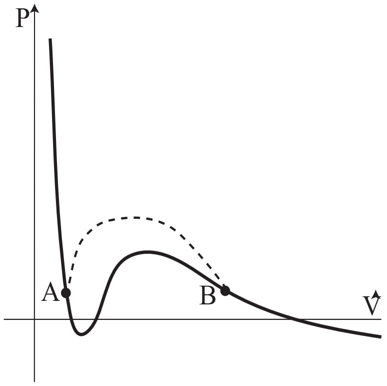
$$E = \frac{3}{2}Nk_BT - \frac{aN^2}{V}. \tag{6.4.35}$$

They are very natural expressions.

**4.4** [Thermodynamically respectable derivation of Maxwell's rule]
If the temperature is sufficiently low, the $PV$-curve given by the van der Waals equation of state implies

$$\frac{\partial P}{\partial V} = -\frac{Nk_BT}{(V-N\sigma)^2} + a\frac{N^2}{V^3} > 0. \tag{6.4.36}$$

That is, it is thermodynamically unrealizable. Actually, gas-liquid coexistence occurs when this 'unphysical behavior' happens, and the coexistence temperature $T$ is determined by the Maxwell rule. This is what Maxwell proposed and an 'explanation' was given in the text but was with a remark that the argument is an abuse of thermodynamics. Many textbooks argue that Maxwell's rule cannot be derived thermodynamically properly, because the argument in the text (the usual one) utilizes thermodynamics where the states are unstable. However, it is possible to avoid this abuse and still we can thermodynamically demonstrate Maxwell's rule. The coexistence condition for phase A and phase B is the agreement of $P$, $T$ and $\mu$. $\mu_B(T,P) - \mu_A(T,P)$ of the difference of the Gibbs free energy must be computable along the path in the phase diagram through only stable phases (that is, the broken curve in the following figure).



Since

$$G = E - ST + PV, \tag{6.4.37}$$

if we compute $E_B - E_A$ and $S_B - S_A$, then $G_A = G_B$ allows us to compute the difference of $PV$, that is, $P(V_A - V_B)$.
(1) Compute $E_B - E_A$.
(2) Compute $S_B - S_A$.
(3) Since $G_B - G_A = 0$, these results allow us to compute $P(V_B - V_A)$. Confirm that this and the result obtained by the naive abuse of thermodynamics:

$$\int_A^B PdV \tag{6.4.38}$$

agree.

**Solution**
(1) We compute the internal energy difference as

$$E_B - E_A = \int_A^B dE = \int_A^B \left[ \left.\frac{\partial E}{\partial V}\right|_T dV + \left.\frac{\partial E}{\partial T}\right|_V dT \right], \tag{6.4.39}$$

where the temperatures at A and at B are identical. To compute the second term we need the constant volume specific heat $C_V$. $C_V$ for the van der Waals gas is identical with that for a (monatomic) ideal gas (the effect of the attractive interaction depends only on the density, so it does not contribute to $C_V$ as can explicitly be seen from (6.4.35)). Since $T_A = T_B$, the second term is always zero. From $dE = TdS - PdV$ we get

$$\left.\frac{\partial E}{\partial V}\right|_T = T\left.\frac{\partial S}{\partial V}\right|_T - P = T\left.\frac{\partial P}{\partial T}\right|_V - P = \frac{1}{2}a\left(\frac{N}{V}\right)^2. \tag{6.4.40}$$

Therefore,

$$E_B - E_A = \int_A^B \frac{1}{2}a\left(\frac{N}{V}\right)^2 dV = \frac{aN^2}{2}\left(\frac{1}{V_A} - \frac{1}{V_B}\right). \tag{6.4.41}$$

(2) This can also be obtained, if we note $T_A = T_B$. The temperature derivative gives $(C_V/T)dT$, but this is a function of $T$ only, so if the initial and the final temperatures are the same, it cannot contribute to the integral. Therefore,

$$S_B - S_A = \int_A^B dS = \int_A^B \left.\frac{\partial S}{\partial V}\right|_T dV. \tag{6.4.42}$$

Thus, we get

$$T(S_B - S_A) = T\int_A^B \left.\frac{\partial P}{\partial T}\right|_V dV = \int_A^B \frac{Nk_BT}{V - N\sigma}dV = Nk_BT\log\frac{V_B - N\sigma}{V_A - N\sigma}. \tag{6.4.43}$$

(3) Since $G_B - G_A = 0$ and since the initial and the final $T$ and $P$ are the same,

$$P(V_B - V_A) = T(S_B - S_A) - (E_B - E_A). \tag{6.4.44}$$

If we introduce (6.4.43) and (6.4.41) into this, along the broken curve in the above figure

$$P(V_B - V_A) = Nk_BT\log\frac{V_B - N\sigma}{V_A - N\sigma} - \frac{aN^2}{2}\left(\frac{1}{V_A} - \frac{1}{V_B}\right). \tag{6.4.45}$$

The RHS of this formula agrees exactly with the naive computation of (6.4.38) along the van der Waals curve. Therefore, Maxwell's rule has been justified thermodynamically. Notice that this happy consequence depends on a peculiar feature of the van der Waals gas that its specific heat is not volume dependent at all.

**4.5** [Grand canonical approach to 1D van der Waals gas]
Let us study the 1D Kac model with the aid of the grand canonical approach.
(1) If there are $N$ particles in the container of volume $V$, the canonical partition function reads

$$Z_N(V) = \int_{(N-1)\sigma}^{V-\sigma} dx_N \cdots \int_{\sigma}^{x_3-\sigma} dx_2 \int_0^{x_2-\sigma} dx_1 \int dp_1\cdots dp_n e^{-\sum_{i=1}^N p_i^2/2mk_BT + aN^2/k_BTV}. \tag{6.4.46}$$

After checking the formula is correct, actually compute this.
(2) Using the result of (1) write down the grand partition function (you cannot perform the summation in a closed form).
(3) The grand partition function written down in (2) has the following structure:

$$\Xi = \sum_{N=0}^{M} e^{VA(N/V)}, \tag{6.4.47}$$

where $M$ is the maximum number of particles we can push into the volume $V$. Show that if the temperature is sufficiently high, there is only one $n = N/V$ that maximizes $A(n)$. Also demonstrate that if the temperature is sufficiently low, there can be three extrema for $A(n)$.
(4) What do you expect the grand partition function looks like, if $n$ that maximizes $A(n)$ are not unique?
(5) There is a text book which writes explicitly as follows:

$$\Xi = e^{\beta PV} + e^{\beta P'V}. \tag{6.4.48}$$

Here, we have assumed that $A(n)$ have two maxima, and the two terms correspond respectively to the two maxima. Is this correct?

**Solution**
(1) Let us start with the $N = 2$ case:

$$Z_2(V) = \frac{1}{h^2}\int_{\sigma}^{V-\sigma} dx_2 \int_0^{x_2-\sigma} dx_1 \int dp_1 dp_2 e^{-\sum_{i=1}^{2} p_i^2/2mk_BT + aN^2/2k_BTV} \tag{6.4.49}$$

$$= \left(\frac{2\pi m k_B T}{h^2}\right)^{2/2} \int_0^{V-2\sigma} dy_2 \int_0^{y_2} dy_1 e^{2aVk_BT} \tag{6.4.50}$$

$$= \left(\frac{2\pi m k_B T}{h^2}\right)^{2/2} \frac{1}{2}(V-2\sigma)^2 e^{2aVk_BT}. \tag{6.4.51}$$

In the above calculation the interparticle distances $y_2 = x_2 - x_1$ and $y_1 = x_1 - 0$ have been introduced.

For $N = 3$, introducing $y_3 = x_3 - x_2$ as well we get

$$Z_3(V) = \frac{1}{h^3}\int_{2\sigma}^{V-\sigma} dx_3 \int_{\sigma}^{x_3-\sigma} dx_2 \int_0^{x_2-\sigma} dx_1 \int dp_1 dp_2 dp_3 e^{-\sum_{i=1}^{3} p_i^2/2mk_BT + aN^2/k_BTV} \tag{6.4.52}$$

$$= \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \int_0^{V-3\sigma} dy_3 \int_0^{y_3} dy_2 \int_0^{y_2} dy_1 e^{9a/2Vk_BT} \tag{6.4.53}$$

$$= \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \frac{1}{3!}(V-3\sigma)^3 e^{9a/2Vk_BT}. \tag{6.4.54}$$

Now, it is easy to guess the following general formula:

$$Z_N(V) = \frac{1}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{N/2} (V - N\sigma)^N e^{aN^2/2Vk_BT}. \tag{6.4.55}$$

It is not hard to show that this is correct in 1D if you plot possible trajectories of particles as a function of time.
(2)

$$\Xi = \sum_{N=0}^{M} \frac{1}{N!}(V-N\sigma)^N (2\pi m k_B T/h^2)^{N/2} e^{aN^2/2Vk_BT} e^{\mu N/k_BT}. \tag{6.4.56}$$

(3) From (6.4.56) we obtain

$$A(N) = N\log(V-N\sigma) - N\log N + N + \frac{N}{2}\log(2\pi m k_B T/h^2) + aN^2/2Vk_BT + \mu N/k_BT, \tag{6.4.57}$$

so

$$A(n) = V[n\log(1/n-\sigma) + n + \frac{n}{2}\log(2\pi m k_B T/h^2) + an^2/2k_BT + \mu n/k_BT] \tag{6.4.58}$$

$$= V[n\log(1/n-\sigma) + n\Lambda + an^2/2k_BT], \tag{6.4.59}$$

where $\Lambda = 1 + (1/2)\log(2\pi m k_B T/h^2) + \mu/k_BT$. Differentiating this wrt to $n$, we get the condition for a maximum:

$$\frac{\partial A(n)}{\partial n} = V\left[\log\left(\frac{1}{n}-\sigma\right) - \frac{1}{1-n\sigma} + \Lambda + an/k_BT\right] = 0. \tag{6.4.60}$$

The first two terms in the brackets are

$$\log(1/n) + \log(1-n\sigma) - \frac{1}{1-n\sigma}. \tag{6.4.61}$$

This is a monotone decreasing function from $+\infty$ (at $n=0$) to $-\infty$ (at $n = 1/\sigma$, the maximum packing density). Therefore, if $T$ is sufficiently large, then there is only one solution to (6.4.60). Also we see there could be three solutions for this equation

if $T$ is sufficiently small; they correspond to two maxima and one minimum between them of $A(n)$.
(4) At high temperatures there is only one maximum for $A(n)$, so we may use the maximum term to estimate (6.4.56). This is just as we have seen in the proof of ensemble equivalence in the text.

$$\Xi \simeq e^{PV/k_BT}. \tag{6.4.62}$$

If $T$ is sufficiently low (with an appropriate chemical potential), as noted in (3) there are two maxima. If the heights of these maxima are different, then thanks to the multiplicative $V$ in the exponent of (6.4.56) only one maximum can contribute. Only when these two maxima have exactly the same heights can they both contribute to the grand partition function, and this corresponds to the phase coexistence temperature.
(5) As already explained in (4) this form holds only at the phase transition temperature. At other temperatures one term is overwhelmingly smaller than the other, and around the taller maximum are numerous higher $A(n)$'s than the secondary maximum, so if we do not pay attention to the former, there is no point to keep the secondary maximum contribution. Thus, if we interpret $e^{\beta PV}$ as representative term(s) among the summands we cannot write such a formula. However, if you interpret each term to be the sum below or above some $n$ (intermediate value of the two coexisting phases), you might be allowed to write such a formula symbolically.

**4.6** [Hard sphere fluid]
The virial equation of state for a fluid interacting with 2-body potential reads

$$\frac{P}{nk_BT} = 1 - \frac{2\pi}{3}\beta n \int_0^\infty \phi'(r)g(r)r^3dr. \tag{6.4.63}$$

Using this formula, derive the equation of state for the hard sphere fluid:

$$\frac{P_H}{nk_BT} = 1 + \frac{2\pi}{3}ng(\sigma)\sigma^3, \tag{6.4.64}$$

where $\sigma$ is the diameter of the sphere. Strictly speaking, $g(\sigma)$ is $\lim_{r\searrow\sigma} g(r)$.

**Solution**
Notice that $\phi_H g$ is zero inside the sphere.

$$\begin{aligned} -\frac{2\pi}{3}n\int_0^\infty \beta\phi_H'(r)g(r)r^3dr &= \frac{2\pi}{3}n\int_0^\infty \frac{df(r)}{dr}e^{\beta\phi}g(r)r^3dr && (6.4.65)\\ &= \frac{2\pi}{3}n\int_0^\infty \delta(r-\sigma)e^{\beta\phi}g(r)r^3dr && (6.4.66)\\ &= \frac{2\pi}{3}ng(\sigma)\sigma^3 && (6.4.67)\end{aligned}$$

**4.7** [Internal energy of fluid interacting with binary forces]
(1) Suppose the interactions among spherical particles can be expressed in terms of the two-body interaction potential $\phi(\boldsymbol{r}) = \phi(r)$. Write down its internal energy in terms of $E$ $\phi$, the number density $n$ and the radial distribution function $g(r)$.
(2) Obtain the internal energy of the Kac fluid (i.e., the fluid interacting via the Kac potential). Set $\sigma = 1$.
(3) Using the virial equation of state

$$P/nk_BT = 1 - \frac{2\pi}{3}n\beta \int_0^\infty \phi'(r)g(r)r^3dr, \tag{6.4.68}$$

obtain the augmented van der Waals equation of state for a Kac fluid:

$$\beta P = \beta P_H + \frac{n^2}{2}\beta \int d^3\boldsymbol{r}\,\phi(\boldsymbol{r}). \tag{6.4.69}$$

Notice that the radial distribution function of the Kac fluid is the same as the had sphere fluid with the same number density.[33]

**Solution**
(1) The internal energy is the expectation value of the system Hamiltonian. The expectation value of the kinetic energy is simple:

$$\left\langle \sum p^2/2m \right\rangle = (3/2)Nk_BT. \tag{6.4.70}$$

The expectation value of the potential energy is

$$\langle U \rangle = \left\langle \frac{1}{2}\sum_{i\neq j} \phi(|\boldsymbol{r}_i - \boldsymbol{r}_j|) \right\rangle = \int d\boldsymbol{x}d\boldsymbol{y}\,\phi(|\boldsymbol{x}-\boldsymbol{y}|) \left\langle \frac{1}{2}\sum_{i\neq j} \delta(\boldsymbol{x}-\boldsymbol{r}_i)\delta(\boldsymbol{y}-\boldsymbol{r}_j) \right\rangle \tag{6.4.71}$$

$$= \frac{1}{2}\int d\boldsymbol{x}d\boldsymbol{y}\, n^{(2)}(\boldsymbol{x},\boldsymbol{y})\phi(|\boldsymbol{x}-\boldsymbol{y}|). \tag{6.4.72}$$

Therefore,

$$E = \frac{3}{2}Nk_BT + 2\pi nN \int dr\, r^2\phi(r)g(r). \tag{6.4.73}$$

---

[33] However, we cannot derive the Maxwell's rule requirement.

(2) For the Kac potential case, the hard-core part of the potential does not contribute to the energy, so we have only to consider the attractive part in the $\gamma \to 0$ limit:

$$U = \frac{1}{2}nN4\pi \int_1^\infty \gamma^3 \phi(\gamma r) g(r) r^2 dr = \frac{1}{2}nN4\pi \int_\gamma^\infty \phi(y) g(y/\gamma) y^2 dy. \tag{6.4.74}$$

$g(y/\gamma) \to 1$ for any finite $y$ in the $\gamma \to 0$ limit, so

$$U = \frac{1}{2}nN4\pi \int_0^\infty \phi(y) y^2 dy = \frac{1}{2}Nn\overline{\phi}. \tag{6.4.75}$$

Therefore,

$$E = \frac{3}{2}Nk_BT + \frac{1}{2}nN\overline{\phi}. \tag{6.4.76}$$

(3) consider the contribution of the interaction potential to the pressure$\beta\Delta P$ (we perform the limit $\gamma \to 0$ at a convenient stage of calculation):

$$\begin{aligned} \beta\Delta P &= -\beta\frac{2\pi}{3}n^2 \int_0^\infty \phi'(r) g(r) r^3 dr & (6.4.77)\\ &= -\beta\frac{2\pi}{3}n^2 \int_0^\infty [\phi_H'(r) + \gamma^4 \phi'(\gamma r)] g(r) r^3 dr. & (6.4.78)\end{aligned}$$

We have already computed the contribution of the hard-core portion:

$$-\frac{2\pi}{3}n^2 \int_0^\infty \beta\phi_H'(r) g(r) r^3 dr = \frac{2\pi}{3}n^2 g(\sigma)\sigma^3. \tag{6.4.79}$$

The contribution from the foothill of the potential is

$$\begin{aligned} -\beta\frac{2\pi}{3}n^2 \int_1^\infty \gamma^4 \phi'(\gamma r) g(r) r^3 dr &= -\beta\frac{2\pi}{3}n^2 \int_\gamma^\infty \phi'(y) g(y/\gamma) y^3 dy & (6.4.80)\\ &\to -\beta\frac{2\pi}{3}n^2 \int_0^\infty \phi'(y) y^3 dy, & (6.4.81)\end{aligned}$$

where we have used $g \to 1$ at infinity. Therefore,

$$\beta\Delta P = +\frac{2\pi}{3}n^2 g(\sigma)\sigma^3 + 2\pi n^2 \beta \int_0^\infty \phi(y) y^2 dy. \tag{6.4.82}$$

**4.8** [Functional differentiation]

Compute the following functional derivative.

$$\frac{\delta}{\delta\psi(\boldsymbol{x})}\int d^3\boldsymbol{r}\left[\frac{1}{2}(\nabla\psi(\boldsymbol{r}))^2-\frac{1}{2}\tau\psi^2(\boldsymbol{r})+\frac{1}{4}\psi^4(\boldsymbol{r})\right], \tag{6.4.83}$$

where $\tau$ is a constant.
(2) [Green's function and functional differentiation] Consider a differential equation (partial or ordinary) $L\varphi = f$, where $L$ is a linear differential operator acting on the functions of $x$.[34] Show that $\delta\varphi/\delta f$ is the Green's function for the initial boundary value problem defined by the linear operator $L$. (This is a problem immediately solved by inspection, but the fact is not meaningless. The method of Green's function is actually a method to solve a differential equation by the first order functional Taylor expansion approach explained in the text; in this case the problem is linear, so the method gives an exact solution.)
(3) Regard the entropy $S$ of a fluid interacting with the binary potential $\phi$ as its functional, and express the functional derivative

$$\frac{\delta S}{\delta\phi(\boldsymbol{r})} \tag{6.4.84}$$

in terms of the radial distribution function (and its appropriate partial derivatives). It may be easy to compute the corresponding functional derivative of the Helmholtz free energy.

**Solution**
(1)

$$\begin{aligned}&\frac{\delta}{\delta\psi(\boldsymbol{x})}\int d^3\boldsymbol{r}\left[\frac{1}{2}(\nabla\psi(\boldsymbol{r}))^2-\frac{1}{2}r\psi^2(\boldsymbol{r})+\frac{1}{4}\psi^4(\boldsymbol{r})\right]\\ =&\int d^3\boldsymbol{r}\left[(\nabla\psi(\boldsymbol{r}))\nabla\delta(\boldsymbol{r}-\boldsymbol{x})-r\psi(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{x})+\psi^3(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{x})\right]=-\nabla^2\psi(\boldsymbol{x})-r\psi(\boldsymbol{x})+\psi^3(\boldsymbol{x})\end{aligned} \tag{6.4.85}$$

(2) The first order Taylor expansion approximation reads

$$\delta\varphi=\int d\bullet\,\frac{\delta\varphi}{\delta f(\bullet)}\delta f(\bullet), \tag{6.4.86}$$

---

[34] If $(Lf)(x)$ is determined by the value of $f$ at $x$ and the values of various derivatives of $f$ at $x$, $L$ is called a differential operator.

where dummy variables are expressed by $\bullet$. The integration wrt the dummy is written as $\int$. If the problem is linear, $\varphi$ does not depend on $f$, so 'integrating' the above formula, we find

$$\varphi = \int d\bullet \, \frac{\delta\varphi}{\delta f(\bullet)} f(\bullet) \tag{6.4.87}$$

Functionally differentiating $L\varphi = f$ wrt $f$ (recall that we may always exchange the order of functional calculus procedures and ordinary calculus procedures), we obtain (in the following $L$ is written as $L_x$ to emphasize that it acts on the functions of $x$)

$$L_x \frac{\delta\varphi(x)}{\delta f(y)} = \delta(x-y) \tag{6.4.88}$$

(under the same linear auxiliary conditions), Hence, we have $G(x|y) = \delta\varphi(x)/\delta f(y)$. Indeed,

$$\begin{aligned} L_x \int G(x|y) f(y) dy &= L_x \int d\bullet \, \frac{\delta\varphi(x)}{\delta f(\bullet)} f(\bullet) = \int d\bullet \, \frac{\delta(L\varphi)(x)}{\delta f(\bullet)} f(\bullet) \quad (6.4.89) \\ &= \int d\bullet \, \frac{\delta f(x)}{\delta f(\bullet)} f(\bullet) = \int d\bullet \, \delta(x-\bullet) f(\bullet) = f(x). \quad (6.4.90) \end{aligned}$$

(3) Let us write $d\Gamma_N = (1/h^{3N}N!)d(\text{phase volume})$ and differentiate the free energy first:

$$\begin{aligned} \frac{\delta A}{\delta\phi(\boldsymbol{r})} &= \frac{\delta}{\delta\phi(r)}(-k_BT \log \int d\Gamma_N \, e^{-\beta(K+\sum\phi(x_i-x_j))} \quad (6.4.91) \\ &= \frac{1}{Z}\int d\Gamma_N \sum_{i<j} \delta(r-(x_i-x_j)) e^{-\beta(K+\sum\phi(x_i-x_j))} = \left\langle \sum_{i<j} d(r-(x_i-x_j)) \right\rangle \quad (6.4.92) \\ &= \frac{1}{2}\left\langle \sum_{i\neq j} \int dy \, \delta(r+y-x_i)\delta(y-x_j) \right\rangle = \frac{1}{2}\int dy \, n^{(2)}(r+y,y). \quad (6.4.93) \end{aligned}$$

Since the system is expected to be translationally symmetric and isotropic, we can simplify this as

$$\frac{\delta A}{\delta\phi(\boldsymbol{r})} = \frac{1}{2}Vn^{(2)}(\boldsymbol{r},0) = \frac{1}{2}Nng(|\boldsymbol{r}|). \tag{6.4.94}$$

Incidentally, this formula gives a perturbative way to compute $A$.
Since $S = -\partial A/\partial T$ (notice that $n$ is a constant)

$$\frac{\delta S}{\delta\phi(\boldsymbol{r})} = -\frac{1}{2}Nn\frac{\partial}{\partial T}g(|\boldsymbol{r}|). \tag{6.4.95}$$

**4.9** [Functional derivatives of canonical partition function]
Redo the calculations (4.4.26)-(4.4.28) using the canonical formalism; this is slightly easier than the grand canonical approach given in the text.

**4.10** [Debye-Hückel theory]
Let us perform the functional Taylor approximation approach explained in Section 4.4 with $A = n$ and $B = U$.
(1) Within this approach find the equation governing the radial distribution function $g$.
(2) Obtain the Fourier transform of the indirect correlation function in the present approximation.
(3) Let $\phi$ be the Coulomb potential. Its Fourier transform may be written as $\phi_k = Q/k^2$. What is the functional form of the indirect correlation function $h(r)$?

**Solution**
(1) From the formulas in the text,we can almost write down the solution as

$$h = -\beta(nh * \phi + \phi). \tag{6.4.96}$$

First, (4.4.31) tells us

$$\left(\frac{\delta n(x|U)}{\delta -\beta U(y)}\right)_{U=0} = n^2 h(x-y) + n\delta(x-y), \tag{6.4.97}$$

so the Taylor approximation reads

$$n(x|\phi) - n(x|0) = \int dy [n^2 h(x-y) + n\delta(x-y)](-\beta\phi(y) - 0). \tag{6.4.98}$$

That is,

$$nh(x) = -\beta n^2 \int dy h(x-y)\phi(y) - \beta n\phi(x). \tag{6.4.99}$$

This leads to the answer given above.
(2) Fourier transforming (6.4.96), we get

$$h_k = \frac{-\beta\phi_k}{1 + \beta n\phi_k}. \tag{6.4.100}$$

(3) For the Coulomb potential

$$h_k = \frac{-\beta Q}{k^2 + \beta n Q}. \tag{6.4.101}$$

Therefore,

$$h \propto \frac{1}{r} e^{-r\sqrt{\beta n Q}}. \tag{6.4.102}$$

This is a Yukawa potential. That is, the Coulomb force is shielded by many-body effects. Recall that $\log g \sim h$ is the effective two-body interaction.

**4.11** [Toy integral equation for liquid]
Let us make a prototype closure for the radial distribution.
(1) Make the first order approximation (i.e., the first order functional Taylor expansion approximation) of $n^{(1)}(\boldsymbol{x}|U)$ in terms of the Boltzmann factor $\exp[-\beta U(\boldsymbol{x})]$.
(2) What is the direct correlation function?
(3) Find the Fourier transform of the indirect correlation function.
(4) Find the equation of state with our approximation with the aid of compressibility or its reciprocal. Assume that the diameter of the hard core is $a$.

**Solution**
Let $A = n^{(1)}(\boldsymbol{x}|U)$ and $B = \exp[-\beta U(\boldsymbol{x})]$. We need the following calculation:

$$\frac{\delta A(\boldsymbol{x}|U)}{\delta B(\boldsymbol{y}|U)} = -e^{\beta U(\boldsymbol{x})} \frac{\delta n^{(1)}(\boldsymbol{x}|U)}{\delta \beta U(\boldsymbol{y})} \tag{6.4.103}$$

$$= e^{\beta U(\boldsymbol{x})}[n^{(2)}(\boldsymbol{x},\boldsymbol{y}|U) - n^{(1)}(\boldsymbol{x}|U)n^{(1)}(\boldsymbol{y}|U) + n^{(1)}(\boldsymbol{x}|U)\delta(\boldsymbol{x}-\boldsymbol{y})]. \tag{6.4.104}$$

Let us estimate this at $U = 0$:

$$\left.\frac{\delta A(\boldsymbol{x}|U)}{\delta B(\boldsymbol{y}|U)}\right|_{U=0} = n^2 g(\boldsymbol{x}-\boldsymbol{y}) - n^2 + n\delta(\boldsymbol{x}-\boldsymbol{y}). \tag{6.4.105}$$

Let $\phi$ be the potential created by the particle placed at the origin. Then, $A(\boldsymbol{x}|\phi) = n^{(1)}(\boldsymbol{x}|\phi) = ng(\boldsymbol{x})$, $A(\boldsymbol{x}|0) = n^{(1)}(\boldsymbol{x}|0) = n$, $B(\boldsymbol{x}|\phi) = \exp[-\beta\phi(\boldsymbol{x})]$ and $B(\boldsymbol{x}|0) = 1$. Using these formulas, the linear Taylor approximation reads

$$ng(\boldsymbol{x}) - n = n\int d\boldsymbol{y}[ng(\boldsymbol{x}-\boldsymbol{y}) - n + \delta(\boldsymbol{x}-\boldsymbol{y})]\left(\exp[-\beta\phi(\boldsymbol{y})] - 1\right). \tag{6.4.106}$$

Using $h$ and Mayer's $f$, we have

$$h(\boldsymbol{x}) = \int d\boldsymbol{y}\,[nh(\boldsymbol{x}-\boldsymbol{y}) + \delta(\boldsymbol{x}-\boldsymbol{y})]f(\boldsymbol{y}), \tag{6.4.107}$$

or

$$h(\boldsymbol{x}) = f(\boldsymbol{x}) + n\int d\boldsymbol{y}\,h(\boldsymbol{x}-\boldsymbol{y})f(\boldsymbol{y}). \tag{6.4.108}$$

This is our final result. This is already a closed equation for $h$.
(2) Comparing (6.4.108) and the Ornstein-Zernike equation, we immediately see $c = f$.
(3) (6.4.108) is linear in $h$, so we can solve it with the aid of Fourier transformation:

$$f(\boldsymbol{k}) = \int d^3\boldsymbol{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} f(\boldsymbol{r}). \tag{6.4.109}$$

Its inverse is

$$f(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int d^3\boldsymbol{k} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} f(\boldsymbol{k}). \tag{6.4.110}$$

The most important feature we use is that the convolution is converted into product:

$$(f * g)(\boldsymbol{k}) = f(\boldsymbol{k})g(\boldsymbol{k}). \tag{6.4.111}$$

Using this, we have

$$h(\boldsymbol{k}) = f(\boldsymbol{k}) + nh(\boldsymbol{k})f(\boldsymbol{k}). \tag{6.4.112}$$

That is,

$$h(\boldsymbol{k}) = f(\boldsymbol{k})/(1 - nf(\boldsymbol{k})). \tag{6.4.113}$$

(4) The compressibility equation gives $\partial P/\partial n$:

$$\beta \left.\frac{\partial P}{\partial n}\right|_{T,V} = 1 - n\int c d\boldsymbol{x} = 1 + n\frac{4\pi}{3}a^3. \tag{6.4.114}$$

Therefore,

$$\frac{P}{nk_BT} = 1 + \frac{2\pi}{3}a^3 n. \tag{6.4.115}$$

**4.12**. [*Scaled particle theory* of hard core fluid][35]
As we know well by now, for a spherical hard core fluid

$$P/nk_BT = 1 + \frac{2\pi}{3}n\sigma^3 g(\sigma), \tag{6.4.116}$$

where $\sigma$ is the diameter of the spherical core. Therefore, to know the hard core equation of state we need $g$ only at $r = \sigma$.

Let $p_0(r)$ be the probability of observing a bubble of radius $r$. Let $nG(r)$ be the expected number of the centers just outside the bubble ($nGdr$ is the expected

---

[35] R. J. Baxter, in *Physical Chemistry an advanced treatise* volume VIIIA Liquid State (edited by H. Eyring, D. Henderson and W. Jost, Academic Press 1971) Chapter 4, Section VIII.

number of particle centers in the spherical shell between $r$ and $r + dr$). When the bubble is of radius $\sigma$, it just behaves as the exclusion zone by the hard sphere at the origin. Therefore,

$$g(\sigma) = G(\sigma). \tag{6.4.117}$$

We have only to determine $G$ to know the hard-core fluid pressure.

(1) Derive

$$p_0(r + dr) = p_0(r)[1 - 4\pi n r^2 G(r)dr]. \tag{6.4.118}$$

That is,

$$\frac{d}{dr}\log p_0(r) = -4\pi n r^2 G(r). \tag{6.4.119}$$

(2) We can determine $G$ for very large $r$. According to the fluctuation theory, the probability of fluctuation that creates a bubble of radius $r$ may be written in terms of the reversible work $W(r)$ required to make it. Therefore,

$$p_0(r) = e^{-\beta W(r)}. \tag{6.4.120}$$

Using this and the thermodynamic result for large $r$ (i.e., for the usual macroscopic bubble!)

$$dW(r) = PdV + f[1 - (2\delta/r)]dA, \tag{6.4.121}$$

where $A$ is the surface area of the bubble, and $f(1 - 2\delta/r)$ is the surface tension of the curved surface of mean curvature $1/r$. Using (6.4.119)-(6.4.121), find $G(r)$ as a function of $r$.

(3) If $r < \sigma/2$, only 1 particle center can come in the bubble. What is this probability? This must be $1 - p_0(r)$ for $r < \sigma/2$.

(4) Determine $G(r)$ for $r < \sigma/2$.

(5) Unfortunately, $G(r)$ is not a smooth function, but it is known that it is continuously differentiable at $r = \sigma/2$. Let us make an approximation that the unknown parameters $f$ and $\delta$ may be determined by matching $G$ obtained in (2) and in (4) at $r = \sigma/2$ smoothly (match $G$ and $G'$ there). Derive, under this approximation,

$$\frac{P}{nk_BT} = \frac{1 + \eta + \eta^3}{(1 - \eta)^3}, \tag{6.4.122}$$

where $\eta$ is the packing density: $\eta = \pi\sigma^3 n/6$ as usual.

[This is identical to PY-C! Furthermore, $f$ obtained is quite reasonable.]

**Solution**

(1) If we assume that there is no more than a single particle in a thin shell, then $4\pi r^2 dr \times nG(r)$ is the expectation value of the number of particles in the shell immediately outside the bubble of radius $r$. Poisson distribution tells us that the

probability to find no particle in the shell is $e^{-4\pi nr^2G(r)dr} = 1 - 4\pi nr^2G(r)dr$.
(2)

$$-\beta\frac{d}{dr}W(r) = -4\pi nr^2G(r). \tag{6.4.123}$$

$A$ and $V$ must be expressed in terms of $r$: $V = (4\pi/3)r^3$, $A = 4\pi r^2$. Therefore,

$$\frac{dW}{dr} = 4\pi Pr^2 + 8\pi f[1-(2\delta/r)]r = 4\pi k_BTnr^2G(r). \tag{6.4.124}$$

That is,

$$G(r) = \frac{P + (2f/r) - (4f\delta/r^2)}{nk_BT}. \tag{6.4.125}$$

(3) $(4\pi/3)nr^3$ is the expectation value of the number of (the centers of the) particles. Here, the sphere of radius $r$ does not contain any particle (with probability $p_0$) or contains only one particle (with probability $p_1 = 1 - p_0$), and the expectation value of the number of particle is equal to $p_1$. Therefore,

$$p_0(r) = 1 - \frac{4}{3}\pi nr^3. \tag{6.4.126}$$

(4) If we use $p_0$ obtained in (3) in (6.4.119), we get

$$G(r) = \frac{1}{1-(4\pi/3)r^3n}. \tag{6.4.127}$$

(5) If we demand $G$ just obtained and (6.4.125) agree at $r = \sigma/2$:

$$\frac{1}{1-(4\pi/3)(\sigma/2)^3n} = \frac{P + (4f/\sigma) - (16f\delta/\sigma^2)}{nk_BT}. \tag{6.4.128}$$

Introducing the following variable

$$X = P/nk_BT,\ \ Y = 4f/n\sigma k_BT,\ \ Z = 4\delta/\sigma, \tag{6.4.129}$$

the above equality reads

$$\frac{1}{1-\eta} = X + Y - YZ. \tag{6.4.130}$$

If we further demand the agreement of $G'$ there, we have

$$\frac{4\pi r^2n}{(1-(4\pi/3)r^3n)^2} = \frac{-(2f/r^2) + 2(4f\delta/r^3)}{nk_BT}, \tag{6.4.131}$$

or

$$\frac{4\pi r^3 n}{(1-(4\pi/3)r^3 n)^2} = \frac{-(2f/r)+2(4f\delta/r^2)}{nk_BT}. \tag{6.4.132}$$

This can be rewritten as follows:

$$\frac{3\eta}{(1-\eta)^2} = -Y + 2YZ. \tag{6.4.133}$$

We need one more relation. It is the expression of the pressure in terms of $G(\sigma)$:

$$X = 1 + 4\eta G(\sigma) = 1 + 4\eta(X + Y/2 - YZ/4), \tag{6.4.134}$$

where the second equality is obtained with the aid of (6.4.125).

From (6.4.130) and (6.4.133) we get

$$YZ = \frac{1}{1-\eta} + \frac{3\eta}{(1-\eta)^2} - X, \tag{6.4.135}$$

$$Y = \frac{2}{1-\eta} + \frac{3\eta}{(1-\eta)^2} - 2X. \tag{6.4.136}$$

Introducing these into (6.4.134) gives

$$X = 1 + 4\eta X + 2\eta Y - \eta YZ, \tag{6.4.137}$$

$$= 1 + 4\eta X + \frac{3\eta}{1-\eta} + \frac{3\eta^2}{(1-\eta)^2} - 3\eta X. \tag{6.4.138}$$

Therefore,

$$(1-\eta)X = 1 + \frac{3\eta}{1-\eta} + \frac{3\eta^2}{(1-\eta)^2} = \frac{1+\eta+\eta^2}{(1-\eta)^2}. \tag{6.4.139}$$

We've done it!

**4.13** [Quantum effect on the second virial coefficient][36]

The second virial coefficient for a spherical symmetrical particle is, classically,

$$B = 2\pi \int_0^\infty (1 - e^{-\beta\phi(r)}) r^2 dr. \tag{6.4.140}$$

Its quantum version should be obtained by replacing the Boltzmann factor $e^{-\beta\phi}$ with the diagonal element of the 2-body density operator $\rho(r)$:

$$B = 2\pi \int_0^\infty (1 - \mathcal{N}\rho(r)) r^2 dr, \tag{6.4.141}$$

---

[36] This is based on T. Kihara, *Molecular forces.*

where $\mathcal{N}$ is the normalization constant to make $\mathcal{N}\rho(r) \to 1$ in the $r \to \infty$ limit.[37]

Thus, the calculation has been reduced to that of $\rho$. Let us perform this through small steps. The difference of fermions and bosons can be ignored except for helium below 25 K. Therefore, we totally ignore the effect of particle symmetry on the wave function. The following calculation was done for the first time by E. Wigner in the 1930s. For $^4$He, below 50 K the quantum correction increases the classical value by about 50%, but by about 10% around 100 K, so the quantum correction is not very small.

(1) Let us assume that two-body Hamiltonian to be

$$H = -\frac{\hbar^2}{2m}(\Delta_1 + \Delta_2) + \phi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|), \tag{6.4.142}$$

where $\Delta_i$ is the Laplacian wrt the position coordinates of the $i$-particle. The two-body density operator is given by $\rho = e^{-\beta H}$. Show that the matrix element $\langle \boldsymbol{r}_1, \boldsymbol{r}_2|\rho|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle$ of the 2-body density operator satisfies the following equation:

$$\frac{\partial}{\partial\beta}\langle \boldsymbol{r}_1, \boldsymbol{r}_2|\rho|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle = -H\langle \boldsymbol{r}_1, \boldsymbol{r}_2|\rho|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle. \tag{6.4.143}$$

If we use the normalized eigenket $|i\rangle$ ($H|i\rangle = E_i|i\rangle$) of $H$, we can write

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2|\rho|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle = \sum_i \langle \boldsymbol{r}_1, \boldsymbol{r}_2|i\rangle e^{-\beta E_i}\langle i|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle. \tag{6.4.144}$$

Here, the summation may include integration.

(2) For the case with $\phi = 0$ (i.e., for the ideal gas case) obtain $\langle \boldsymbol{r}_1, \boldsymbol{r}_2|\rho|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle$. As already noted, you can totally forget about the particle exchange symmetry. You must specify the initial and the boundary conditions correctly to solve the above parabolic equation.

(3) Let us introduce the deviation $\Psi$ from the case without interactions as

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2|\rho|\boldsymbol{r}_1', \boldsymbol{r}_2'\rangle = \left(\frac{2\pi\hbar^2\beta}{m}\right)^{-3} \exp\left[-\frac{m}{2\hbar^2\beta}[(\boldsymbol{r}_1 - \boldsymbol{r}_1')^2 + [(\boldsymbol{r}_2 - \boldsymbol{r}_2')^2] + \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2; \boldsymbol{r}_1', \boldsymbol{r}_2')\right]. \tag{6.4.145}$$

Note that $-k_BT\Psi$ corresponds to the 'quantum-corrected two-body interaction. $\Psi$ satisfies the following equation:

$$\phi + \frac{\partial\Psi}{\partial\beta} + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}_1')\cdot\nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}_2')\cdot\nabla_2]\Psi = \frac{\hbar^2}{2m}[(\nabla_1\Psi)^2 + \Delta_1\Psi + (\nabla_2\Psi)^2 + \Delta_2\Psi]. \tag{6.4.146}$$

---

[37] As we will see below, the diagonal element of the position representation of the density operator is proportional to the probability of finding particles, so it is proportional to the Boltzmann factor semi-classically. Therefore, normalizing the density operator appropriately, we can interpret it as a quantum statistical extension of the spatial Boltzmann factor.

This equation can be obtained from (6.4.143) by substituting (6.4.145); nothing special has not been done at all.

To obtain the quantum correction we expand $\Psi$ as

$$\Psi = \Psi_0 + \frac{\hbar^2}{2m}\Psi_1 + \cdots \tag{6.4.147}$$

and then introduce this into (6.4.146). Requiring the order by order agreement of the substituted result, we get

$$\phi + \frac{\partial}{\partial\beta}\Psi_0 + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}'_1)\cdot\nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}'_2)\cdot\nabla_2]\Psi_0 = 0, \tag{6.4.148}$$

$$\frac{\partial}{\partial\beta}\Psi_1 + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}'_1)\cdot\nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}'_2)\cdot\nabla_2]\Psi_1 = (\nabla_1\Psi_0)^2 + \Delta_1\Psi_0 + (\nabla_2\Psi_0)^2 + \Delta_2\Psi_0, \tag{6.4.149}$$

etc. First, we must solve the zeroth order equation. Show that the diagonal element of the only meaningful solution is $\Psi_0(\beta, \boldsymbol{r}_1, \boldsymbol{r}_2; \boldsymbol{r}_1, \boldsymbol{r}_2) = -\beta\phi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$.

(4) We have only to obtain the diagonal element of $\Psi_1$. Solve the simplified equation that can be obtained by taking the diagonal limit of (6.4.149):

$$\frac{\partial}{\partial\beta}\Psi_1 = \lim_{\{\boldsymbol{r}'_i\}\to\{\boldsymbol{r}_i\}}\left[(\nabla_1\Psi_0)^2 + \Delta_1\Psi_0 + (\nabla_2\Psi_0)^2 + \Delta_2\Psi_0\right]. \tag{6.4.150}$$

As can immediately be seen, to compute the RHS of this equation, we cannot use $\Psi_0(\beta, \boldsymbol{r}_1, \boldsymbol{r}_2; \boldsymbol{r}_1, \boldsymbol{r}_2) = -\beta\phi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$ that is already in the diagonal limit. The derivative must be computed from the original expression of $\Psi_0$. Find $\lim_{\{\boldsymbol{r}'_i\}\to\{\boldsymbol{r}_i\}}\nabla_1\Psi_0$ and $\lim_{\{\boldsymbol{r}'_i\}\to\{\boldsymbol{r}_i\}}\nabla_2\Psi_0$.

(5) To obtain $\Delta_1\Psi_0$ we could perform a similar calculation. However, our purpose is not to practice calculation (although this practice is also meaningful), the needed result is provided:

$$\Delta_i\Psi_0 = -\frac{1}{3}\beta\Delta_i\phi, \tag{6.4.151}$$

where $i = 1, 2$.

Obtain the diagonal element $\langle\boldsymbol{r}_1, \boldsymbol{r}_2|\Psi_1|\boldsymbol{r}_1, \boldsymbol{r}_2\rangle$. Then compute the quantum correction to the second virial coefficient to order $\hbar^2$. Show that the correction is positive for any $T$ (as stated for helium, quantum correction always increases the second virial coefficient).

(6) The quantum correction starts with the term of $O[\hbar^2]$. Why, or why is't there any correction of odd order in $h$?

**Solution**

(1) Obviously,

$$\frac{d}{d\beta}e^{-\beta H} = -He^{-\beta H}. \tag{6.4.152}$$

If we position represent this, we get the desired equation. The result is a parabolic partial differential equation.
(2) We must solve the following partial differential equation:

$$\frac{d}{d\beta}f(\beta,\{\boldsymbol{r}_i\},\{\boldsymbol{r}'_i\}) = \frac{\hbar^2}{2m}\left(\Delta_1+\Delta_2\right)f(\beta,\{\boldsymbol{r}_i\},\{\boldsymbol{r}'_i\}), \tag{6.4.153}$$

where the primed vector variables are regarded as constant parameters in this equation. The initial condition at $\beta = 0$ can be obtained immediately from the definition of the density matrix:

$$f(0,\{\boldsymbol{r}_i\},\{\boldsymbol{r}'_i\}) = \sum_i \langle \boldsymbol{r}_1,\boldsymbol{r}_2|i\rangle\langle i|\boldsymbol{r}'_1,\boldsymbol{r}'_2\rangle = \langle \boldsymbol{r}_1,\boldsymbol{r}_2|\boldsymbol{r}'_1,\boldsymbol{r}'_2\rangle = \delta(\boldsymbol{r}_1-\boldsymbol{r}'_1)\delta(\boldsymbol{r}_2-\boldsymbol{r}'_2). \tag{6.4.154}$$

The boundary condition is that the solution goes to zero if $\{\boldsymbol{r}_i\}$ and $\{\boldsymbol{r}'_i\}$ are far apart.

This equation can be totally separated for different particles (as can be guessed from physics), so we have only to solve

$$\frac{\partial}{\partial\beta}\varphi(\beta,\boldsymbol{r},\boldsymbol{r}') = \frac{\hbar^2}{2m}\Delta\varphi(\beta,\boldsymbol{r},\boldsymbol{r}'). \tag{6.4.155}$$

This is an elementary diffusion equation (the simplest way to solve it is to use Fourier transformation):

$$\varphi(\beta,\boldsymbol{r},\boldsymbol{r}') = \left(\frac{2\pi\hbar^2\beta}{m}\right)^{-3/2}\exp\left[-\frac{m}{2\hbar^2\beta}(\boldsymbol{r}-\boldsymbol{r}')^2\right]. \tag{6.4.156}$$

Therefore, the answer is (6.4.145) without $\Psi$. We immediately see that $\mathcal{N}$ can be chosen to cancel the numerical factor in front of the exponential function.
(3) The characteristic equation of this linear first order partial differential equation is

$$\frac{d\beta}{1} = -\frac{d\Psi_0}{\phi} = \frac{\beta dx_1}{x_1-x'_1} = \frac{\beta dy_1}{y_1-y'_1} = \cdots, \tag{6.4.157}$$

where $\cdots$ denote similar equations for al the remaining particles. From this we see that the diagonal element is $\Psi_0 = -\beta\phi +$ const. The answer must agree with the

ideal case in th $\beta\phi \to 0$ limit, the constant must be zero.
(4) If we apply $\nabla_1$ to the zeroth order equation (6.4.148), we get

$$\nabla_1\phi + \frac{\partial}{\partial\beta}\nabla_1\Psi_0 + \frac{1}{\beta}\nabla_1\Psi_0 + k_BT[(\boldsymbol{r}_1 - \boldsymbol{r}_1')\cdot\nabla_1 + (\boldsymbol{r}_2 - \boldsymbol{r}_2')\cdot\nabla_2]\nabla_1\Psi_0 = 0. \quad (6.4.158)$$

Its diagonal element is

$$\nabla_1\phi + \frac{\partial}{\partial\beta}\nabla_1\Psi_0 + \frac{1}{\beta}\nabla_1\Psi_0 = 0. \quad (6.4.159)$$

This is an ordinary differential equation, so its general solution reads

$$\nabla_1\Psi_0 = -\frac{1}{2}\beta\nabla_1\phi + k_BTc, \quad (6.4.160)$$

where $c$ is an arbitrary function of $\boldsymbol{r}$. However, we know that both $\Psi$ and $\Psi_0$ must vanish in the $\beta \to 0$ limit, $c = 0$. Similarly,

$$\nabla_2\Psi_0 = -\frac{1}{2}\beta\nabla_2\phi. \quad (6.4.161)$$

(5) Since

$$\langle\boldsymbol{r}_1, \boldsymbol{r}_2|\Psi_1|\boldsymbol{r}_1, \boldsymbol{r}_2\rangle = \frac{\beta^3}{6}(\nabla\phi)^2 - \frac{\beta^2}{3}\Delta\phi, \quad (6.4.162)$$

we now know the 'quantum-corrected' two-body potential. Therefore, we can write

$$B(T) = \frac{1}{2}\int_V d\boldsymbol{r}(1 - e^{\Psi}) = B_c(T) - \frac{\hbar^2}{4m}\int_V\left(\frac{\beta^3}{6}(\nabla\phi)^2 - \frac{\beta^2}{3}\Delta\phi\right)e^{-\beta\phi}d^3\boldsymbol{r}, \quad (6.4.163)$$

where $B_c$ is the classic value. If we perform integration by parts, the result becomes cleaner. Using Gauss' theorem, we have

$$(\Delta\phi)e^{-\beta\phi} = \nabla\cdot((\nabla\phi)e^{-\beta\phi}) - \nabla\phi\cdot\nabla e^{-\beta\phi}, \quad (6.4.164)$$

so

$$\int(\Delta\phi)e^{-\beta\phi} = \int_S e^{-\beta\phi}\nabla\phi\cdot d\boldsymbol{S} - \int d^3\boldsymbol{r}\nabla\phi\cdot\nabla e^{-\beta\phi} = +\beta\int d^3\boldsymbol{r}(\nabla\phi\cdot\nabla\phi)e^{-\beta\phi}. \quad (6.4.165)$$

Thus, we have arrived at the desired result:

$$B(T) = B_c(T) + \frac{\hbar^2}{4m}\int_V\frac{\beta^3}{6}(\nabla\phi)^2 e^{-\beta\phi}d^3\boldsymbol{r}. \quad (6.4.166)$$

Thanks to the spherical symmetry $\nabla\phi(r) = (\boldsymbol{r}/r)\phi'(r)$ (the chain rule!), we can further rewrite this as

$$B(T) = B_c(T) + \frac{\pi\beta^3\hbar^2}{6m}\int_0^\infty (\phi')^2 e^{-\beta\phi} r^2 dr. \tag{6.4.167}$$

The correction term is clearly positive for any $T$.
(6) Perform $h \to -h$ in quantum mechanics. For the system with spatial inversion symmetry, this corresponds to the time reversal operation. We are discussing equilibrium states, so the situation should not be affected by time reversal. Therefore, quantum correction must be even in $h$. [What if the system is chiral?]

## 6.5 Problems for Chapter 5

**5.1** [Phase transition and analyticity]
If there is no phase transition in a range of fugacity $z = e^{\beta\mu}$, $P/k_BT$ is, by definition, holomorphic in $z$, so we may expand it as

$$\frac{P}{k_BT} = \sum_{\ell=1}^{\infty} b_\ell z^\ell, \tag{6.5.1}$$

where $b_\ell$ is called the $\ell$-body cluster coefficient. They are smooth (actually real analytic) functions of $T$ and positive for lower temperatures, so each $b_\ell$ has the smallest real positive zero $T_\ell$. It is known that $\{T_\ell\}$ is a monotone decreasing sequence of $\ell$. It is demonstrated[38] that
(i) $b_\ell(T_c) > 0$ if $\ell$ is sufficiently large.
(ii) There are infinitely many $T_\ell$ between any $T$ $(> T_c)$ and $T_c$.
Let $T_a$ be the accumulation point of $\{T_\ell\}$. Show $T_c = T_a$.[39]

**Solution**
(ii) implies that $T_c$ is an accumulation point of $\{T_\ell\}$. Needless to say, $\{T_\ell\}$ is bounded from below and monotone decreasing, its accumulation point is unique. Therefore, $T_c = T_a$. (i) is unnecessary.

**5.2** [Crude version of rigorous Peierls' argument]
Let us impose an all up spin boundary condition to the 2-Ising model on the finite square. Then, we wish to take a thermodynamic limit. If the spin at the center of the square is more likely to be up than to be down, we may conclude that there is a long-range order.

Let $\gamma$ be a closed Bloch wall (i.e., the boundary between up and down spin domains; this does not mean that the domain enclosed by $\gamma$ is a totally up or down domain (lakes can contain islands with ponds with islets, etc.; the wall corresponds to the shore lines.) The probability $P_V(\gamma)$ to find such a wall in the system with volume $V$ has the following estimate (we used this in our discussion on Peierls' argument):

$$P_V(\gamma) \le e^{-2\beta J|\gamma|}, \tag{6.5.2}$$

where $|\gamma|$ is the total length of the contour $\gamma$, $\beta = 1/k_BT$, and $J$ is the usual ferromagnetic coupling constant. [This naturally looking inequality needs a proof; it

---

[38] T. Kihara and J. Okutani, Chem. Phys. Lett., **8**, 63 (1971).

[39] This problem asks a mathematically trivial question, but the main point is the fact stated here.

is not trivial.]
(1) Since the outside boundary is all up, there must be a Bloch wall encircling the origin for the spin at the origin to be down. Therefore, the probability $P_V^0$ of the spin at the origin to be down must be smaller than the probability of the occurrence of at least one contour encircling the origin. Show

$$P_V^0 \le \sum_{\gamma} e^{-2\beta J|\gamma|}, \tag{6.5.3}$$

where the summation is over all the possible contours surrounding the origin. [Hint: Don't think too much. If event $A$ occurs when at least one of $B$ and $C$ occurs, then $A \subset B \cup C$.]
(2) Estimate the number of contours with circumference $|\gamma|$ crudely as the total number of random walks of length $|\gamma|$ starting from appropriate neighborhood points of the origin. Use this crude estimate and show that if $\beta$ is sufficiently large, $P_V^0 < 1/2$ for large $V$.

**Solution**
The argument here is, although simplified, almost rigorous.
(1) The event that the spin at the origin is down occurs only if at least one Bloch wall surrounds the origin. Let $P(\gamma)$ be the probability that there is a closed Bloch wall $\gamma$ surrounding the origin. Then,

$$P_V^0 \le P(\text{at least there is one Bloch wall surrounding the origin}) = P(\cup\{\exists\gamma\}) \le \sum_{\gamma} P(\gamma), \tag{6.5.4}$$

where $\cup\{\exists\gamma\}$ is the event that there is at least one closed wall around the origin irrespective of its shape. The left-most inequality is due to $P(A\cup B) \le P(A) + P(B)$. The sum is over all the closed curves around the origin. We use (6.5.2) to get (6.5.3).
(2) Again we perform extremely excessive estimate: to draw a closed curve surrounding the origin of length $|\gamma|$, we must start at some point. Let the starting point be the closest point on $\gamma$ to the origin. The number of candidate points for this cannot exceed $|\gamma|^2$. Let us start a random walk of length $|\gamma|$. There is no guarantee that the walk makes a closed curve, but all the curves satisfying the desired condition can be drawn in this fashion. Since the number of distinct walks is $4 \times 3^{|\gamma|-1}$, we have

$$P_V^0 \le \sum_{\gamma} e^{-2\beta J|\gamma|} \le \sum_{|\gamma|=4}^{\infty} |\gamma|^2 3^{|\gamma|} e^{-2\beta J|\gamma|}. \tag{6.5.5}$$

Here, we have used the fact that the smallest closed curve on the lattice surrounding the origin has length 4 (the smallest square on the dual lattice; notice that the wall is a curve on the dual lattice). Therefore, if $\beta$ is sufficiently large,for any $V$ we can make $P_V^0 < 1/2$. The key point of the argument is that $[\log C_{|\gamma|}]/|\gamma|$ is uniformly bounded from above, where $C_n$ is the total number of closed curves of length $n$ surrounding the origin.

(6.5.5) tells us that the spin at the origin points upward more likely than downward. The required $\beta$ is unrealistically large due to the crudeness of the estimation, but still it is finite (that is, $T > 0$). Thus, we have shown that the system orders at some low but positive temperature.

Now, we demonstrate (6.5.2).[40] This is called Peierls' inequality. The statement of the proposition is:

Let $V$ be a finite domain. On its boundary all the spins are fixed to point upward. Under this condition the probability $P_V(\gamma)$ of formation of a Bloch wall $\gamma$ is bounded from above as

$$P_V(\gamma) \le e^{-\beta 2J|\gamma|}, \tag{6.5.6}$$

where $|\gamma|$ is the total length of $\gamma$.

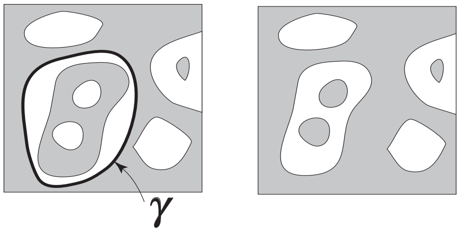
To demonstrate this, the energy in the volume $V$ is estimated (let us denote the number of lattice points in $V$ by the same symbol $V$):

$$\mathcal{H}_V(\phi) = -2JV + 2J|\partial\phi(V)|, \tag{6.5.7}$$

where $\partial\phi$ denotes the totality of the Bloch walls appearing in the spin configuration $\phi$. We may write

$$P_V(\gamma) = \frac{\sum_{\phi(V):\gamma\subset\partial\phi(V)} e^{-\beta\mathcal{H}_V(\phi)}}{\sum_{\phi(V)} e^{-\beta\mathcal{H}_V(\phi)}}. \tag{6.5.8}$$

Here, in the numerator the sum is over all the spin configurations containing the Bloch wall $\gamma$. Next, the spin configurations on $V$ is divided into the ones including $\gamma$ denoted by $\Phi_\gamma$ and the rest denoted by $\Phi_\gamma^-$. Let us define a one-to-one map $\chi : \Phi_\gamma \to \Phi_\gamma^-$ that flips all the spins inside $\gamma$ (see the figure below).



Configurations with and without $\gamma$. This correspondence is one-to-one.

---

[40] Chapter 2 Section 1 of Ya G Sinai, *Theory of Phase Transitions: Rigorous Results* (Pergamon Press, 1982).

The map $\chi$ reduce the number of adjacent up-down spin pairs by $|\gamma|$, so the Boltzmann factor must be multiplied by $e^{2\beta|\gamma|}$. Therefore, if we replace every element in $\Phi_\gamma$ with the corresponding element in $\Phi_\gamma^-$, we must multiply $e^{-2\beta|\gamma|}$ to cancel this Boltzmann factor:

$$\begin{aligned} P_V^+(\gamma) &= \frac{\sum_{\Phi_\gamma} e^{-\beta\mathcal{H}_V(\phi)}}{\sum_{\phi(V)} e^{-\beta\mathcal{H}_V(\phi)}}. && (6.5.9) \\ &= e^{-2\beta|\gamma|}\frac{\sum_{\Phi_\gamma^-} e^{-\beta\mathcal{H}_V(\phi)}}{\sum_{\phi(V)} e^{-\beta\mathcal{H}_V(\phi)}} \le e^{-2\beta J|\gamma|}. && (6.5.10) \end{aligned}$$

**5.3** [Phase transition in 1D long-range system]
Using Peierls' argument, discuss the phase transition in a 1d spin system whose coupling constant behaves as $r^{-q}$ ($q < 2$) beyond some distance $r_0$ (you may assume that the coupling constant for $r < r_0$ is $J$, constant). No rigorous argument is wanted.

**Solution**
Assume initially all the spins are up. Then, we flip $L$ contiguous spins. The required energy is estimated as

$$\Delta E = 2\int_{L/2+\delta}^{\infty} dx \int_{-L/2}^{L/2} dy \frac{1}{(x-y)^q}. \tag{6.5.11}$$

The contribution of the short-range interactions is $\sim 2J\delta$, so it is ignored. It is easy to see

$$\int_L^\infty dx \int_{-L/2}^{L/2} dy \frac{1}{(x-y)^q} = O[L^{2-q}]. \tag{6.5.12}$$

This energy increases indefinitely with $L$ if $q < 2$. As we have seen for the nearest neighbor interaction system in the text, the contribution of entropy that fatally cripple system ordering is of order $\log L$, so the entropy effect cannot destroy the order even at finite temperatures.

According to a rigorous argument even for $q = 2$ phase transition occurs at a finite temperature.

**5.4** [Griffiths' inequality]
Empirically, it is known that there is the following relation among critical exponents:

$$\alpha + \beta(1+\delta) = 2. \tag{6.5.13}$$

(1) Thermodynamically, demonstrate the following inequality (Griffiths' inequality)

$$\alpha + \beta(1+\delta) \geq 2. \tag{6.5.14}$$

[Hint: You may proceed just as the case of Rushbrooke's inequality, but use $m \sim h^{1/\delta}$ (at $T = T_c$) to differentiate wrt $H$ under constant temperature. At $\tau = 0$ and $h = 0$ the (vertical) line $T = T_c$ is tangent to $m = m(\tau, h = 0)$, so $m$ may be parameterized by $\tau$.]
(2) Using the scaling relation or Kadanoff's argument (that is, using the expression of the magnetization as a generalized homogeneous function of $h$ and $\tau$), show that the equality actually holds. [Hint: the wisest approach may be to use $\beta\delta = \beta + 1$. You may use such as $\alpha + 2\beta + \gamma = 2$ we have already discussed.]

**Solution**
(1) We start with the following inequality (just as in the case of Rushbrooke's inequality):

$$\left.\frac{\partial S}{\partial T}\right|_H \left.\frac{\partial M}{\partial H}\right|_T \geq \left.\frac{\partial S}{\partial H}\right|_T \left.\frac{\partial M}{\partial T}\right|_H = \left.\frac{\partial M}{\partial T}\right|_H^2. \tag{6.5.15}$$

Using $m \sim h^{1/\delta}$ at $T_c$, very close to $T_c$ we have

$$T^{-1}|\tau|^{-\alpha}h^{1/\delta-1} = |\tau|^{-\alpha}m^{1-\delta} \geq |\tau|^{2(\beta-1)}, \tag{6.5.16}$$

where the $T$ in the LHS is almost $T_c$, so we may ignore it. Using the hint, we have

$$|\tau|^{-\alpha}|\tau|^{\beta(1-\delta)} \geq |\tau|^{2(\beta-1)}. \tag{6.5.17}$$

This requires that the LHS of the following formula must be larger than some constant:

$$|\tau|^{-(\alpha+\beta(1+\delta)-2)} \geq \text{ const}, \tag{6.5.18}$$

which implies $\alpha + \beta(1+\delta) - 2 \geq 0$.
(2) Using the following result we have obtained in the text:

$$\beta = (d - y_2)/y_1, \delta = y_2/(d - y_2), \gamma = (2y_2 - d)/y_1, \tag{6.5.19}$$

we get

$$\beta\delta = y_2/y_1 = \beta + \gamma. \tag{6.5.20}$$

This with $\alpha + 2\beta + \gamma = 2$ gives what we want.
[In Griffiths' original (PRL **14** 623 (1965)), primed and unprimed quantities are undistinguished. This distinction corresponds the values below and above $T_c$. Now, we believe this distinction is not needed, so in the present exposition, primes were

removed.]

**5.5** [A toy illustration of Lee-Yang theory][41]
Suppose the grand partition function of a fluid in a volume $V$ is given by

$$\Xi_V = (1+z)^V \frac{1-z^V}{1-z}, \tag{6.5.21}$$

where $z$ is the fugacity.
(1) Find the zeros of $\Xi_V$. How does the distribution of the zeros change as $V \to \infty$?
(2) Obtain $P$ for real positive $z$ and locate the phase transition.
(3) Find the volume $v$ per particle as a function of $z$.
(4) Find the $P$-$v$ relation, and demonstrate that phase coexistence does happen.

**Solution**
(1) Zeros are $-1$ and $e^{-2\pi ik/V}$ $(k = 1, 2, \cdots, V)$. Therefore, as $V \to \infty$, the density of zeros on the unit circle increases. In particular on the real positive axis, $z = 1$ is an accumulation point of zeros. Therefore, $z = 1$ must be a phase transition point.
(2) Let us take the thermodynamic limit. If $|z| < 1$,

$$\beta P_V = \frac{1}{V}\log\left[(1+z)^V \frac{1-z^V}{1-z}\right] \to \log(1+z). \tag{6.5.22}$$

If $|z| > 1$,

$$\beta P_V = \frac{1}{V}\log\left[(1+z)^V \frac{1-z^V}{1-z}\right] = \frac{1}{V}\log\left[(1+z)^V z^V \frac{z^{-V}-1}{1-z}\right] \to \log z + \log(1+z). \tag{6.5.23}$$

Therefore,

$$\beta P = \begin{cases} \log(1+z) & z < 1, \\ \log z + \log(1+z) & z > 1. \end{cases} \tag{6.5.24}$$

As expected, $z = 1$ is a singularity and phase transition occurs there.
(3) Since

$$\frac{\partial \beta P_V}{\partial \log z} = \frac{z}{1+z} - \frac{z^V}{1-z^V} + \frac{1}{V}\frac{z}{1-z}, \tag{6.5.25}$$

we have

$$\frac{1}{v} = \lim_{V\to\infty} \frac{\partial \beta P_V}{\partial \log z} = \begin{cases} z/(1+z) & z < 1, \\ (2z+1)/(1+z) & z > 1. \end{cases} \tag{6.5.26}$$

---

[41] S similar question can be found in Reichl

Notice that for $z > 1$ and for $z < 1$ the number densities (or the specific volumes) are distinct; of course, the smaller $z$ corresponds to gas.
(4) For $z < 1$, we have $v = 1 + 1/z$, so $v > 2$, and

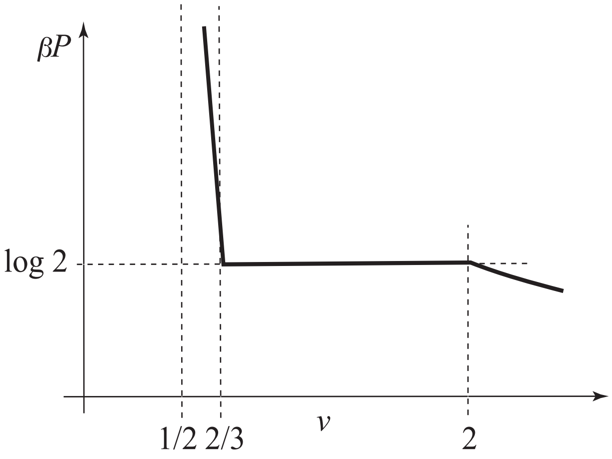
$$\beta P = \log \frac{v}{v-1}. \tag{6.5.27}$$

Notice that this is equal to $\log 2$ at $v = 2$.

For $z > 1$ since $z = (1-v)/(2v-1)$, $1/2 < v < 2/3$ and

$$\beta P = \log \frac{v(1-v)}{(2v-1)^2}. \tag{6.5.28}$$

This gives $\log 2$ at $v = 2/3$. Therefore, in the interval $v \in [2/3, 2]$ the pressure is constant: $\beta P = \log 2$; we expect a phase coexistence. The equation of state can be plotted as



**5.6** [The Lee-Yang circle theorem illustrated
The theorem is proved in Supplementary Pages. Here, let us check its content for simple cases.
(1) According to the theorem the root of $p(z) = 1 + 2az + z^2$ must be on the unit circle as long as $a \in [-1, 1]$. Confirm this.
(2) On the apices of a triangle are spins interacting with each other. For this system construct the polynomial of $z$, and confirm that as long as the interactions are ferromagnetic, all the zeros are on the unit circle.

**Solution**

(1) This is the circle theorem for the two spin system. The roots can be computed

explicitly and $-a \pm i\sqrt{1-a^2}$, so they are on the unit circle.
(2) The polynomial in the theorem reads

$$p(z) = 1 + 3az + 3az^2 + z^3, \tag{6.5.29}$$

where $a = e^{-2\beta J}$ (even if the interactions are different for different spin pairs, the modification of the following argument is easy and no different argument is needed), which is less than 1 for ferromagnetic interactions. This polynomial is a real positive coefficient polynomial, so according to the theorem $-1$ must be a root. It is indeed the case: we can factorize the polynomial as

$$p(z) = (z+1)(z^2 + (3a-1)z + 1). \tag{6.5.30}$$

The complex roots are

$$\frac{1}{2}\left\{(1-3a) \pm i\sqrt{4-(1-3a)^2}\right\}, \tag{6.5.31}$$

which are actually on the unit circle. Very interestingly, as soon as the intereaction ceases to be ferromagnetic, the theorem breaks down.

**5.7** [A derivation of mean field theory]
A mean field approach may be obtained with the aid of a variational principle for free energy. If the (density) distribution function of microstates is $f$ (we consider classical case) the Helmholtz free energy may be written as:

$$A = \langle H \rangle + k_BT \int d\Gamma f(\Gamma) \log f(\Gamma). \tag{6.5.32}$$

Here, the integration is over the whole phase space. $\langle H \rangle$ is the expectation value of the system Hamiltonian with respect to $f$. Let us apply this to the Ising model on a $N \times N$ square lattice. Its Hamiltonian is as usual

$$H = -J \sum_{\langle i,j \rangle} S_i S_j. \tag{6.5.33}$$

If we could vary $f$ unconditionally and minimize $A$, then the minimum must be the correct free energy, but this is in many cases extremely hard or plainly impossible. Therefore, we assume an approximate form for $f$ and the range of variation is narrowed. For example, we could introduce a ‘single-body’ approximation:

$$f = \phi(S_1)\phi(S_2)\cdots\phi(S_N), \tag{6.5.34}$$

where $\phi$ is a single-spin (density) distribution function.
(1) Under this approximation write down $A$ in terms of $\phi$. That is, find $X_1$ and $X_2$ in the following formula:

$$A = X_1 \left[ \sum_S \phi(S) S \right]^2 + X_2 \sum_S \phi(S) \log \phi(S), \tag{6.5.35}$$

where $N$ and $N \pm 1$ need not be distinguished.
(2) Minimize $A$ wrt $\phi$. $\phi$ must be normalized. What is the equation determining $\phi$?
(3) Using the obtained formula, write down the magnetization per spin. The used Lagrange's multiplier must be determined.

**Solution**
(1) We have only to compute each term honestly:

$$\langle H \rangle = -J \sum_{\langle i,j \rangle} \sum_{S_k = \pm 1} S_i S_j \prod_k \phi(S_k) \tag{6.5.36}$$

$$= -J \sum_{\langle i,j \rangle} \langle S_i \rangle \langle S_j \rangle = -2JN^2 \langle S \rangle^2. \tag{6.5.37}$$

Hence, $X_1 = 2JN$.

The entropy part reads

$$k_B T \sum_{S_i = \pm 1} \prod_k \phi(S_k) \sum_k \log \phi(S_k) = N^2 k_B T \sum_{s = \pm 1} \phi(s) \log \phi(s). \tag{6.5.38}$$

Therefore, $X_2 = N^2 k_B T$.
(2) Introducing a Lagrange's multiplier $\lambda$ to impose the normalization condition, we must minimize

$$A + \lambda \sum_{S=\pm 1} \phi(S) = -2N^2 J \left[ \sum_{S=\pm 1} \phi(S) S \right]^2 + N^2 k_B T \sum_{S=\pm 1} \phi(S) \log \phi(S) + \lambda \sum_{S=\pm 1} \phi(S). \tag{6.5.39}$$

The minimization condition reads

$$-4N^2 J S \langle S \rangle + N^2 k_B T (1 + \log \phi(S)) + \lambda = 0, \tag{6.5.40}$$

so we see

$$\phi(S) \propto \exp(-4\beta J m S), \tag{6.5.41}$$

where $\langle S \rangle = m$.
(3) We can get rid of $\lambda$ from the formula for $m$ with the aid of $\langle S \rangle = m$:

$$m = \tanh 4J\beta m. \tag{6.5.42}$$

This is nothing but the self-consistency equation.

**5.8** [Gibbs-Bogoliubov' inequality and mean field]
(1) Derive the following inequality (called the Gibbs-Bogoliubov inequality) with the aid of Jensen's inequality for classical systems [This question has already been asked in Chapter2, but is reproduced here]:

$$A \le A_0 + \langle H - H_0 \rangle_0. \tag{6.5.43}$$

Here, $A$ is the free energy of the system with the Hamiltonian $H$, $A_0$ is the free energy of the system with the Hamiltonian $H_0$, $\langle\ \rangle_0$ is the average over the canonical distribution wrt $H_0$. [Hint: compute $\left\langle e^{-(H-H_0)} \right\rangle_0$; the temperature may be absorbed or we could use the unit system with $k_B = 1$.]

All the variational approximations for statistical thermodynamics are applications of this inequality.[42] Let $H$ be the Hamiltonian of the system we are interested in, and $H_0$ be the Hamiltonian of a system whose free energy $A_0$ we can compute exactly. We introduce variational parameters in $H_0$ and tehn try to make the RHS of (6.5.43) as small as possible.
(2) As $H$ we adopt the $N \times N$ 2-Ising model Hamiltonian (without a magnetic field; even with it there is almost no change), and

$$H_0 = \sum_i h s_i. \tag{6.5.44}$$

Derive the equation for $h$ that minimizes the RHS of (6.5.43).

**Solution**
(1) See the solution to 2.22.
(2) Let us first compute $A_0$ and $m = \langle s_i \rangle_0$:

$$A_0 = -k_B T \log [2 \cosh \beta h]^{N^2}, \tag{6.5.45}$$

$$m = -\tanh \beta h. \tag{6.5.46}$$

---

[42] See, for example, M. D. Girardeau and R. M. Mazur, "Variational methods in statistical mechanics," Adv. Chem. Phys. XXIV, eds. I. Prigogine and S.A. Rice (Academic, New York, 1974), p187-255.

Consequently, since the number of nearest neighbor pairs is $2N^2$,

$$\langle H \rangle_0 = -J(2N^2)\tanh^2 \beta h. \tag{6.5.47}$$

Also we have

$$\langle H_0 \rangle_0 = -N^2 h \tanh \beta h. \tag{6.5.48}$$

Combining all the results, we can write the Gibbs-Bogoliubov inequality as

$$A \le -N^2 k_B T \log[2\cosh \beta h] - J(2N^2)\tanh^2 \beta h + N^2 h \tanh \beta h. \tag{6.5.49}$$

Differentiating the RHS wrt $h$, we have

$$-N^2 \frac{\sinh \beta h}{\cosh \beta h} - 4JN^2 \tanh \beta h \frac{\sinh \beta h}{\cosh^2 \beta h} + N^2 \tanh \beta h + N^2 h \frac{\sinh \beta h}{\cosh^2 \beta h} = 0, \tag{6.5.50}$$

so we obtain

$$4J \tanh \beta h = h. \tag{6.5.51}$$

If we multiply $\beta$ to the both sides and taking their tanh, we get

$$\tanh(4\beta J \tanh \beta h) = \tanh \beta h, \tag{6.5.52}$$

but if we use (6.5.46), this turns out to be our familiar formula:

$$m = \tanh 4\beta J m. \tag{6.5.53}$$

**5.9** [Exact mean field for 1-Ising model]

The starting point of the mean-field theory can be the following exact relation for the 1-Ising model:

$$\langle s_0 \rangle = \langle \tanh \beta J(s_{-1} + s_1) \rangle, \tag{6.5.54}$$

where $\langle\ \rangle$ is the equilibrium expectation. Utilizing $s^2 = 1$ and translational symmetry of the system, write down a closed equation for $m = \langle s \rangle$, and then discuss the possible phase transitions.

**Solution**

The odd powers of $s_1$ and $s_{-1}$ are identical to the original $s_1$ and $s_{-1}$, respectively, we can write

$$\tanh \beta J(s_{-1} + s_1) = A(s_{-1} + s_1), \tag{6.5.55}$$

where $A$ is a numerical constant. If we set $s_1 = s_{-1} = 1$, we have $\tanh 2\beta J = 2A$, which fixes $A$, and we get the following identity:

$$m = (\tanh 2\beta J)\, m. \tag{6.5.56}$$

Unless $\beta = \infty$, $\tanh 2\beta J < 1$, so $m = 0$ for any positive temperature; we know there is no phase transition for $T > 0$. If $\beta \to \infty$, the equation reads $m = m$, so the theory does not exclude the possibility of $T_c = 0$, but it seems we cannot say anything further.

**5.10** [2-Ising model on the honeycomb lattice]
Let us consider a 2-Ising model on the honeycomb lattice whose coupling constant is $J$. Assume there is no magnetic field.
(1) Find the equation corresponding to (5.8.4).
(2) Find $T_c$ with the aid of the approximation corresponding to (5.8.6).
(3) Then, using a more accurate mean field theory corresponding to (5.8.13) compute $T_c$. Which $T_c$ obtained by (2) or this question should be lower? Is your result consistent with your expectation?

**Solution**
(1) The coordination number of the honeycomb lattice is 3, so

$$\langle s_0 \rangle = \langle \tanh[\beta J(s_1 + s_2 + s_3)] \rangle. \tag{6.5.57}$$

(2) The approximation gives

$$m = \tanh 3\beta J m. \tag{6.5.58}$$

That is,

$$x = 3\beta J \tanh x. \tag{6.5.59}$$

This gives $3\beta_c J = 1$ or $T_c = 3J/k_B$.
(3) The equation corresponding to (5.8.10) is

$$\tanh \beta J(s_1 + s_2 + s_3) = a(s_1 + s_2 + s_3) + b s_1 s_2 s_3, \tag{6.5.60}$$

and the coefficients are determined by the following simultaneous equation

$$\tanh 3\beta J = 3a + b, \tag{6.5.61}$$

$$\tan \beta J = a - b. \tag{6.5.62}$$

We get

$$a = \frac{1}{4}(\tanh \beta J + \tanh 3\beta J), \quad b = \frac{1}{4}(\tanh 3\beta J - 3\tanh \beta J). \tag{6.5.63}$$

Thus, the mean-field equation reads

$$m = 3am + bm^3, \tag{6.5.64}$$

or

$$m = \sqrt{\frac{1-3a}{b}}. \tag{6.5.65}$$

This implies $T_c$ is determined by $a = 1/3$:

$$\tanh \beta J + \tanh 3\beta J = 4/3. \tag{6.5.66}$$

A more accurate calculation is expected to take the effect of fluctuations more accurately into account. Fluctuations oppose ordering, so better approximation should give lower $T_c$. That is, we can expect that the $T_c$ from the current approximation method is lower than that obtained in (3), i.e., $T_c = 3J/k_B$.

It is not hard to prove that the $T_c$ according to the 'better' approximation is indeed lower than $3J/k_B$, but here let us use a numerical result: $\beta_c J = .47$ or $T_c = 2.13J/k_B$. The exact answer is known to be $T_c = 1.52J/k_B$; the improvement is considerable.

**5.11** [1-Gaussian model]
At each lattice point $i$ of a one-dimensional lattice lives a real variable $q_i$, and the system Hamiltonian is given by

$$H = \sum_j \left[\frac{1}{2}q_j^2 - Kq_jq_{j+1}\right]. \tag{6.5.67}$$

The partition function reads

$$Z = \left(\prod_j \int_{-\infty}^{\infty} dq_j\right) \prod_{j=1}^{N-1} \exp[w(q_j, q_{j+1})], \tag{6.5.68}$$

where

$$w(x,y) = -\frac{1}{4}(x^2+y^2) + Kxy. \tag{6.5.69}$$

The partition function should be evaluated just as the 1-Ising model with the aid of the eigenvalue problem:

$$\lambda f(x) = \int_{-\infty}^{\infty} dy\, f(y) \exp\left[-\frac{1}{4}(x^2+y^2) + Kxy\right]. \tag{6.5.70}$$

The integral kernel is Gaussian, so the eigenfunction belonging to the largest eigenvalue should be of constant sign [Hint: you can see a correspondence to the transfer matrix approach; actually, there is a counterpart of Perron-Frobenius theorem for positive definite integral kernels]. Therefore, we may assume that $f$ is also Gaussian.

(1) Find $f(x)$ (its multiplicative numerical coefficient may be ignored).
(2) Find the free energy per lattice point. Is there any phase transition? ***y***

**Solution**

(1) Since he integral kernel is $L^2$ (square integrable), the integral operator is compact (a Hilbert-Schmidt operator), so the spectrum is discrete.[43] The integral kernel is positive definite, so the largest eigenvalue is positive and non-degenerate (a counterpart of the Perron-Frobenius theorem holds for compact operators). We are discussing an integral operator, so there are infinitely may eigenvalues, but they accumulate at 0. We have only to consider the eigenvalue corresponding to the Perron-Frobenius eigenvalue. Assume

$$f(x) = e^{-ax^2/2}. \tag{6.5.71}$$

Then, we have

$$\int dy\, e^{-ay^2/2} \exp\left[-\frac{1}{4}(x^2+y^2) + Kxy\right] = \sqrt{\frac{2\pi}{a+1/2}} \exp\left(-\frac{1}{2} + \frac{K^2}{a+1/2}\right) x^2. \tag{6.5.72}$$

If we choose $a$ as

$$a = \frac{1}{2} - \frac{K^2}{a+1/2}, \tag{6.5.73}$$

or

$$a = \sqrt{(1/4) - K^2}, \tag{6.5.74}$$

$f$ becomes an eigenfunction belonging to the following eigenvalue:

$$\lambda = \sqrt{\frac{4\pi}{1+\sqrt{1-4K^2}}} \tag{6.5.75}$$

(2) The problem is quite parallel to the ordinary spin problems; the free energy per lattice is given by

$$f = -\frac{1}{2} k_B T \log \frac{4\pi}{1+\sqrt{1-4K^2}}. \tag{6.5.76}$$

If $f$ is holomorphic wrt $K$, there is no phase transition. A candidate singularity is at $K = \pm 1/2$, and it is a branching point. In the current problem, the system loses stability for $|K| > 1/2$ (the free energy is not bounded from below), so we conclude that there is no phase transition in the range of parameters where the model is meaningful.

---

[43] Kolmogorov-Fomin, *Introductory Functional Analysis* )Dover) is an excellent textbook to review such common-sense knowledge.

**5.12** [Correlation function by mean field theory, or mean field for nonuniform space]
(1) Let us assume that the coupling constant and the magnetic field depend on spatial position:

$$H = -\sum_{\langle i,j\rangle} J_{ij} s_i s_j - \sum_i h_i s_i. \tag{6.5.77}$$

Derive the basic equation for the mean-field theory for a square lattice:

$$\langle s_i \rangle = \left\langle \tanh(\beta \sum_j J_{ij} s_j + \beta h_i) \right\rangle. \tag{6.5.78}$$

If we introduce the crude approximation like (5.8.6), we obtain

$$\langle s_i \rangle = \tanh\left(\beta \sum_j J_{ij} \langle s_j \rangle + \beta h_i\right). \tag{6.5.79}$$

(2) We wish to compute the spatial correlation $\langle s_i s_j \rangle$. First, demonstrate that

$$\frac{\partial \langle s_i \rangle}{\partial h_k} = k_B T \langle s_i s_k \rangle \tag{6.5.80}$$

without any approximation.
(3) Applying this to the following form of (6.5.79), obtain the equation for $\{\langle s_i s_k \rangle\}$:

$$\mathrm{Arctanh} \langle s_i \rangle = \beta \sum_j J_{ij} \langle s_j \rangle + \beta h_i. \tag{6.5.81}$$

(4) Now, let us go over to the continuum limit, assuming that the system has a translational symmetry. If we write the correlation as $g$, the equation obtained in (3) becomes

$$\int dy \left( \frac{\delta(x-y)}{1-m^2} - \beta J(x-y) \right) g(y-z) = \delta(x-z). \tag{6.5.82}$$

We have already assumed the spatial translational symmetry and set $m(x) = m$. The Fourier transform of the coupling constant reads

$$J(k) = \sum_j e^{ik\cdot r_j} J(r_j) = 2^d J \cos k_x \cos k_y \cdots. \tag{6.5.83}$$

Find the Fourier transform $G(k)$ of the correlation function $g$.
(5) If we are interested in global features, we have only to pay attention to small $k$. Determine the coefficients $A$ and $B$ in the following formula (you may assume $T > T_c$):

$$G(k) \simeq \frac{1}{A + Bk^2}. \tag{6.5.84}$$

(6) Determine the critical exponent $\nu$.

**Solution**

(1) no explanation should be needed.
(2) This is an exact relation without any approximation.

$$\begin{aligned}\frac{\partial}{\partial h_k}\langle s_i\rangle &= \frac{\partial}{\partial h_k}\frac{\sum s_i \exp\left(\beta J \sum s_j s_k + \beta \sum h_k s_k\right)}{\sum \exp\left(\beta J \sum s_j s_k + \beta \sum h_k s_k\right)} && (6.5.85)\\ &= \beta\frac{\sum s_i s_k \exp\left(\beta J \sum s_j s_k + \beta \sum h_k s_k\right)}{\sum \exp\left(\beta J \sum s_j s_k + \beta \sum h_k s_k\right)} - \beta\langle s_i\rangle\langle s_k\rangle. && (6.5.86)\end{aligned}$$

(3) Differentiating (6.5.81) with $h_k$, we obtain

$$\beta\frac{1}{1-\langle s_i\rangle^2}\langle s_i s_k\rangle = \beta^2\sum_j J_{ij}\langle s_j s_k\rangle + \beta\delta_{ik}. \tag{6.5.87}$$

That is,

$$\frac{1}{1-\langle s_i\rangle^2}\langle s_i s_k\rangle = \beta\sum_j J_{ij}\langle s_j s_k\rangle + \delta_{ik}. \tag{6.5.88}$$

(4) Fourier transformation is defined as follows:

$$G(k) = \int dx\, e^{ikx} g(x). \tag{6.5.89}$$

(6.5.88) becomes

$$\left(\frac{1}{1-m^2} - \beta J(k)\right) G(k) = 1, \tag{6.5.90}$$

so we obtain

$$G(k) = \frac{1}{1/(1-m^2) - \beta J(k)}. \tag{6.5.91}$$

(5) Since we may assume $T > T_c$, $m = 0$. An approximation of $J(k)$ for small $k$ is obtained from (6.5.83) as

$$J(k) = 2^d J\left(1 - \frac{1}{2}\boldsymbol{k}^2 + \cdots\right), \tag{6.5.92}$$

so

$$G(k) \simeq \frac{1}{1 - 2^d\beta J + 2^{d-1}\beta J \sum k_i^2}. \tag{6.5.93}$$

Therefore,

$$A = 1 - 2^d\beta J, \;\; B = 2^{d-1}\beta J. \tag{6.5.94}$$

(6) The mean-field critical point is determined by $2^d\beta J = 1$ so we may write $A \propto (T - T_c)$. On the other hand $B$ is almost a constant, so taking positive constants $a$ and $b$, the Fourier transform of the correlation function reads

$$G(k) = \frac{1}{a(T - T_c) + bk^2}. \tag{6.5.95}$$

Therefore, the critical exponent for the correlation length is $\nu = 1/2$.

**5.13** [Lattice gas on honeycomb lattice]
Let us relate the 2-Ising model on the honeycomb lattice and the lattice gas on the same lattice. The Ising Hamiltonian $H$ and the lattice gas Hamiltonian $H_L$ as just as given in the text:

$$H = -J\sum_{\langle i,j\rangle} s_is_j - h\sum_i s_i, \tag{6.5.96}$$

$$H_L = -J'\sum_{\langle i,j\rangle} n_in_j. \tag{6.5.97}$$

Let $V$ ($\gg 1$) be the total number of lattice points and down spins are regarded particles.
(1) Following the procedure around p260, rewrite the canonical partition function of the Ising model in therms of the number of down spins $[D]$ and that of down spin pairs $[DD]$.
(2) Express the lattice gas pressure in terms of magnetic field $h$ and the free energy per spin $f$.
(3) Demonstrate that the lattice gas pressure $P$ is a continuous function of $h$.
(4) Sketch the free energy $Vf = -k_BT\log Z$[44] of the Ising model for a few representative temperatures. Next, sketch the pressure of the lattice gas as a function of $\log z$ (this is essentially the chemical potential) for a few representative temperatures. Then, explain their noteworthy features succinctly.

**Solution**
(1) $V = [U] + [D]$ so

$$\sum_i s_i = -[D] + [U] = V - 2[D]. \tag{6.5.98}$$

---

[44] Notice that this free energy is $G$ rather than $A$, since $h$ is on.

The total energy is

$$-J\sum_{\langle i,j\rangle} s_i s_j - h\sum_i s_i = J([UD]-[UU]-[DD]) + h(2[D]-V). \tag{6.5.99}$$

Since

$$2[UU]+[UD]=6[U],\ \ 2[DD]+[UD]=6[D], \tag{6.5.100}$$

we have

$$[UD]-[UU]-[DD]=[UD]-\left(3[U]-\frac{1}{2}[UD]\right)-\left(3[D]-\frac{1}{2}[UD]\right)=2[UD]-3V. \tag{6.5.101}$$

Therefore (subtracting the energy $-J$ from each spin energy),

$$H = 2J[UD]+h(2[D]-V) = -hV+(12J+2h)[D]+4J[DD]. \tag{6.5.102}$$

This equation happens to be identical to (5.3.14). The canonical partition function for the spin system reads

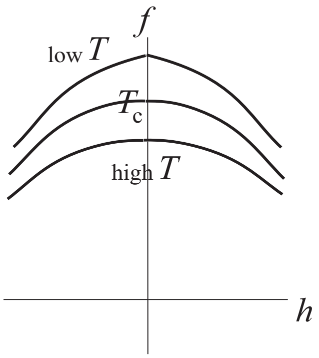
$$Z=\sum_C \exp\{-\beta(-hV+(2h+12J)[D]-4J[DD])\}, \tag{6.5.103}$$

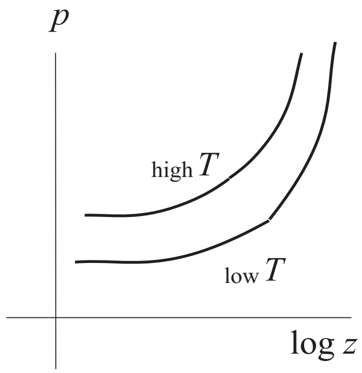
where $\sum_C$ implies the some over all the spin configurations.
(2) since $[D]$ is the total number of particles, $e^{-\beta(2h+12J)}=z$ is the fugacity. $H_L = -4J[DD]$ and

$$Ze^{-\beta hV} = e^{-\beta V(h+f)} = \sum z^{[D]}e^{4\beta J[DD]}. \tag{6.5.104}$$

That is, the Pressure of the lattice gas is given by $P=-h-f$.
(3) We know from the result of the Ising model that $f$ is a continuous function of $h$. Therefore, $P$ is continuous. Even if we lack such knowledge about the Ising model, we may use the logic explained below (5.3.5). Needless to say, the number density is bounded from above. If $V$ is finite, then $P$ is differentiable, and its derivative is uniformly bounded from above. Therefore, $P$ in the thermodynamic limit must be continuous.
(4) If the temperature $T$ is very high, there is no phase transition, so $f$ is a smooth function (actually a holomorphic function) of $h$. If $|h|$ is increased, $f$ decreases. $f$ is an even function of $h$. If we lower $T$, eventually we encounter a second-order phase transition, and then first order phase transitions. This happens for $h=0$, so $f$ starts to have a kink at $h=0$. Thus, the sketch of $f$ is as in the left of the following figure.

As a function of $\log z$ $P=-h-f$ may be guessed easily, since $\log z$ is essentially $-h$. The crucial point is that the $P$-$\log z$ curve exhibits a kink corresponding to the





gas-liquid phase transition.

See Fig. 2 of the first paper by Lee and Yang: Phys. Rev., **87**, 404 (1952).

**5.14** [RG by Migdal approximation[45]]
When we discussed 'decimation', we have realized that the procedure is not very good in the space higher than 1D. For example, if we apply the method to the 2-Ising model (taking $\ell = 2$, i.e., thin half of spins), we obtain

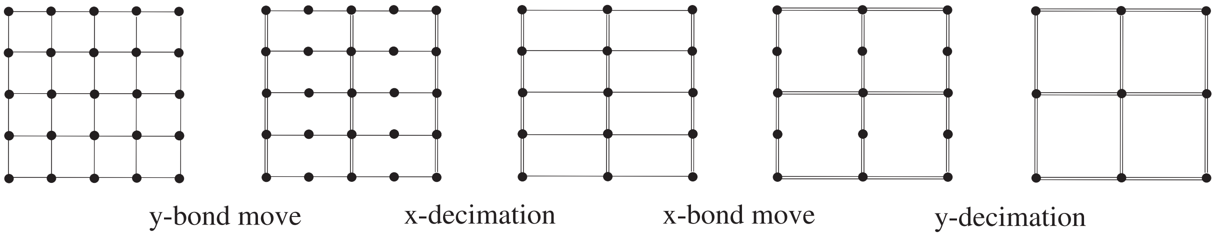
$$K' = \frac{1}{4}\log\cosh 4K, \tag{6.5.105}$$

where the Hamiltonian is written in the following form:

$$H = -\sum K s_i s_j \tag{6.5.106}$$

and the temperature is absorbed in the parameter. The fixed point of this transformation is $K = 0$ (i.e., the high temperature limit), so there is no ordering.

Migdal proposed to remedy the defect of underestimating the interactions as follows (see the figure below).



---

[45]

(i) [$y$-bond moving step] Every other vertical bonds ($y$-bonds) are combined with their right-neighboring bonds . If the coupling constant in the $y$-direction is $K_y$, the coupling constant due to the new bonds made by combining two vertical couplings is $2K_y$.
(ii) [$x$-decimation step] For the $x$-direction, one dimensional thinning is performed. For the new $x$-directional coupling constant is computed by the 1D thinning result we obtained (5.10.10).
(iii) [$x$-bond moving step] Next, every other $x$-bonds are merged with their lower neighbor $x$-bonds.
(iv) [$y$-decimation step] Apply one-dimensional decimation in the $y$-direction.
Thus, we have arrived at the square lattice with the lattice spacings doubled (i.e., $\ell = 2$). If we halve the spatial scale we can complete a renormalization group transformation.
(1) Let us put $'$ to the parameters after the procedure (i)-(iv). Show that

$$K_x' = \log\cosh(2K_x), \tag{6.5.107}$$

$$K_y' = \frac{1}{2}\log\cosh(4K_y). \tag{6.5.108}$$

Here, the 'initial values' are $K$ for both the $x$ and $y$ couplings. Notice that in two different directions, the step-cycle of the procedure is 'out of phase,' so to speak. In the $x$-direction, the decimation is applied first and then the bond are moved, while in the $y$-direction the bonds are merged first, and then decimation follows. Consequently, the fixed points of these two equations have different fixed points (marked with $^*$): $K_x^* = 2K_y^*$.[46]
(2) Find all the fixed points $K_x^*$. Which corresponds to the critical fixed point?
(3) Linearizing the renormalization transformation around the fixed point, we can calculate critical exponents; we have only to compute $d\log\cosh(2K_x)/dK_x$. This corresponds to $\ell^{y_1}$. Determine $\nu$. We cannot say the result is impressive, but still there is an improvement from the mean-field approach.

**Solution**
(1) These formula should be obvious, if you understand the 1D decimation.
(2) 0 and $\infty$ are stable fixed points and they correspond to the high and low temperature limits, respectively. If you draw the graph of the LHS, there is one more fixed point, which is unstable. This must be the most interesting fixed point. This is determined by (1/2)Arc cosh $e^x = x$ (iterative substitution could solve this): we have $K^* = 0.609$.
(3) The derivative is $2\tanh 2K$, which is 1.678 at $K^*$. Therefore, $1.678 = 2^{1/\nu}$ or

---

[46] This might be remedied by changing the ordering every other RG step; that is, with obvious abbreviations: $M_y \to D_x \to M_x \to D_y \to M_x \to D_y \to M_y \to D_x$.

$1/\nu = 0.746$, that is, $\nu = 1.339$. The mean field result is 1/2, and the exact result is 1.

**5.15** [Finite size effect]
The specific heat of a certain magnetic system behaves $C \sim |\tau|^{-\alpha}$ near its critical point without external magnetic field, if the specimen is sufficiently large. If the magnet is not very large, or more concretely, if it is a sphere of radius $R$, near its critical point, its maximum specific heat is $C(R)$. Compute the ratio $C(2R)/C(R)$ in terms of critical indices.

**Solution**
We expect that $C$ would depends on $R$, but the dependence must be only through a dimensionless parameter. The most natural dimensionless quantity near $Tc$ must be its ratio against the correlation length. Thus, $C = \tau^{-\alpha} f(R/\xi)$. Since $\xi \sim \tau^{-\nu}$, we have

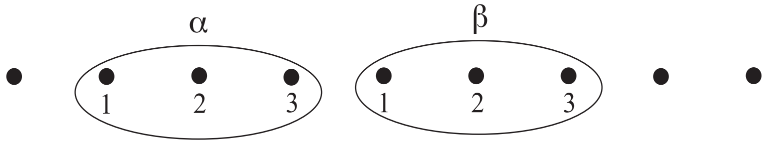
$$C = \tau^{-\alpha} f(\tau^{\nu} R). \tag{6.5.109}$$

If we are sufficiently close to the critical point, the effect of the size (i.e., $R$) becomes prominent, and the temperature effect should be masked: in the $\tau \to 0$ limit $C$ must not depend on $\tau$. To this end, the function $f(x)$ around $x = 0$ must be a power: $f(x) \propto x^q$. $\tau^{-\alpha}(\tau^{\nu})^q \sim 1$ implies $q = \alpha/\nu$. Consequently, we see

$$C \sim R^{\alpha/\nu}. \tag{6.5.110}$$

Hence, $C(2R)/C(R) \simeq 2^{\alpha/\nu}$.

**5.16** [Use of block spins in 1-Ising model]
Let us construct an RG transformation for a 1-Ising model with a similar approach as is applied to the triangle lattice1-Ising model. We start with (5.10.12). The figure corresponding to Fig. 5.10.2 os as shown below:



The equation corresponding to (5.10.13) is

$$K' s'_{\alpha} s'_{\beta} = K s_{\alpha 3} s_{\beta 1}, \tag{6.5.111}$$

and $s'$ is the block spin of ±1 determined by the majority rule. This relation cannot literally be realized, so just as in the triangle lattice case, we need an analogue of (5.10.14).
(1) Write down $\phi(K)$ corresponding to (5.10.16).

(2) Write down the RG equation corresponding to (5.10.19) and (5.10.21).
(3) Find the fixed points.
(4) What can you conclude from these calculations?

**Solution**

(1) We wish to demand that

$$K' s'_\alpha s'_\beta = K s_{\alpha 3} s_{\beta 1} \tag{6.5.112}$$

is an identity, but this is impossible, so we demand that the identity holds on the average. Therefore, we demand

$$K' s'_\alpha s'_\beta = K \langle s_{\alpha 3} \rangle_{s'_\alpha} \langle s_{\beta 1} \rangle_{s'_\beta}, \tag{6.5.113}$$

where $\langle s_{\alpha 3} \rangle_{s'_\alpha}$ is the conditional expectation value of $s_{\alpha 3}$ under the condition $s'_\alpha = 1$ or $-1$. The up-down symmetry of the spin system (for simplicity we consider the case with $h \simeq 0$, so you may assume $h = 0$), so if we understand the $s'_\alpha = 1$ case, the rest is obvious. We can trace what we have done for the triangle lattice. The table corresponding to the one just above (5.10.16) reads

|  | $+++$ | $++-$ | $+-+$ | $-++$ |
|---|---|---|---|---|
| $s_{\alpha 2}$ | $+1$ | $+1$ | $-1$ | $+1$ |
| $s_{\alpha 3}$ | $+1$ | $-1$ | $+1$ | $+1$ |
| $E$ | $-2J$ | $0$ | $+2J$ | $0$ |

(6.5.114)

Following the logic in the text, we have

$$\langle s_{\alpha 3} \rangle_{s'_\alpha} = \frac{e^{2K} + e^{-2K}}{2 + e^{2K} + e^{-2K}} s'_\alpha = \frac{\cosh 2K}{1 + \cosh 2K} s'_\alpha. \tag{6.5.115}$$

$\langle s_{\alpha 1} \rangle_{s'_\alpha}$ is quite the same.

(2) From the above result, we get

$$K' = K \left( \frac{\cosh 2K}{1 + \cosh 2K} \right)^2 . \tag{6.5.116}$$

Actually, we also need $\langle s_{\alpha 2} \rangle_{s'_\alpha}$:

$$\langle s_{\alpha 2} \rangle_{s'_\alpha} = \frac{e^{2K} - e^{-2K} + 2}{2 + e^{2K} + e^{-2K}} = \frac{1 + \sinh 2K}{1 + \cosh 2K}. \tag{6.5.117}$$

This gives

$$h' = h \frac{1 + \sinh 2K + 2\cosh 2K}{1 + \cosh 2K}. \tag{6.5.118}$$

(3) We have to solve

$$K_F = K_F\left(\frac{\cosh 2K_F}{1+\cosh 2K_F}\right)^2. \tag{6.5.119}$$

$K_F = 0$ and $\cosh 2K_F = +\infty$ that is $K_F = +\infty$ are the fixed points. $h_F = 0$ is obvious. Hence, $(0,0)$ is a stable fixed pint and $(+\infty, 0)$ an unstable fixed point.[47]
(4) Obviously, $(0,0)$ corresponds to the high temperature limit. $(\infty, 0)$ corresponds to $T \searrow 0$. Since it is a repeller, we expect that the fixed point corresponds to a critical point at $T = 0$.

**5.17** ['Democracy'][48]
Let us consider a hierarchical organization in which the decision at the $k$th level depends on the decisions of the $s$ cells of the $k-1$st level. Assume that the 0th level corresponds to individual members of the organization, and the decision at level $F$ is the decision of the organization. If there are sufficiently many levels (actually 5 levels are enough), the system may be understood as a system to coarse-grain individual opinions. To be frank, any political organization is a coarse-graining mechanism of opinions, and it is usually the case that conscientious subtle voices do not reach the top.

In the following, we assume there are two options A and B that must be chosen. Consider the fraction $p_n$ of the cells at level $n$ that support B.
(1) Suppose $s = 3$ and strict majority rule is applied. Find all the fixed points of this system and study their stability.
(2) Suppose $s = 4$. Majority rule is applied but if two opinions A and B are equally supported, A is always selected. For B to be the decision of the organization, at least how many % of the people should support B? In the extreme case, if $s = 2$, what happens?
(3) Suppose $s = 5$. Majority rule is applied, but due to the organizational propaganda at every level there is always at least one cell that supports A. For B to win despite this arrangement, what is the minimum % of the supporters of B?

**Solution**
(1) If 3 or 2 cells vote for B, the choice of the cell is B, so if $p_n$ is the fraction of B-supporter at the $n$th level,

$$p_{n+1} = p_n^3 + 3p_n^2(1-p_n). \tag{6.5.120}$$

---

[47] Here, $K_F = -\infty$ is also a fixed point, and this corresponds to the antiferromagnetic case. However, we ignore this, because, if the initial system (the actual system we wish to study) is ferromagnetic and $K > 0$ the RG flow never go into the $K < 0$ region.

[48] cf. S. Galam and S. Wonczk, "Dictatorship from majority rule voting," Eur. Phys. J. B **18**, 183 (2000). The paper contains some trivial calculation errors, so trust your own result.

The fixed point $f$ obeys

$$f = f^3 + 3f^2(1-f), \tag{6.5.121}$$

or

$$f(1-f)(2f-1) = 0. \tag{6.5.122}$$

Therefore, $f = 0, 1$ and $1/2$. 0 and 1 are stable fixed points, and $1/2$ is unstable. This scheme is fair in the sense that the the majority (whose fraction is $1/2 +$ infinitesimal number) regulated the whole organization, and can change the regime. [Really? Of course, voting or democracy can decide only not very serious questions; you cannot nationalize foreign companies only by voting.]

(2) In this case to choose B 4 or 3 cells must choose B:

$$p_{n+1} = p_n^4 + 4p_n^3(1-p_n). \tag{6.5.123}$$

This case is discussed in the quoted paper. The fixed point fractions $f$ obey

$$f(1-f)(1+f-3f^2) = 0, \tag{6.5.124}$$

so 0, 1 and $(1+\sqrt{13})/6 = 0.76759$ are the fixed points. It is very hard (77% support required) to change the regime.

If $s = 2$, then

$$p_{n+1} = p_n^2, \tag{6.5.125}$$

so it is impossible to change the regime.

As you have already realized, the problem is a caricature of constructing block spins or Kadanoff transformation $\mathcal{K}$. What if we choose a coarse-graining procedure that destroys the system Hamiltonian? In this case, automatically the field appears that tries to restore the system symmetry.

(3) This is exactly the case of (2).

# Index