Statistical Physics: a jogging course (2018-9)\textsuperscript{1}

Feb 14, 2019

Yoshi Oono
Physics UIUC
yoono@illinois.edu (3111 ESB)

You can use the following notes as your elementary calculus check list:
Also you can use the following notes as your elementary linear algebra re-introduction:
http://www.yoono.org/Y_OONO_official_site/P427_S19_lectures_files/P427LinearAlgebra.pdf

I conceived a bit different introductory course from the traditional Phys 427; I wish to connect elementary gas kinetics and Brownian motion smoothly to equilibrium statistical thermodynamics. Thus, as is clear, this course emphasizes three levels of description of the world, microscopic, mesoscopic and macroscopic descriptions. It is also emphasized that the latter two are closely related to large deviation and the law of large numbers, respectively.

Needless to say, it is not very easy to cover these topics within one semester, so the course is a jogging course. As an undergrad course in the US it is a bit challenging. However, every student graduating from physics should know at least superficially the topics covered in these lectures. If 1/3 of the participating students think interesting and rewarding, the course is a success.\textsuperscript{2} Thus, when these lectures were

\textsuperscript{1}While the previous version was, upon the suggestion of Cambridge University Press, expanded into a full year course for advanced undergrads (perhaps, one semester for beginning graduates) and published from the same publisher as Perspectives on Statistical Thermodynamics, this is a cleaned version with some updates of the original jogging course. I agreed with Cambridge that I can continue to use the shorter (and older) lecture version for my own course. You do not need the book; do not buy it for this course.

\textsuperscript{2}V. Arnold said, “M. Faraday arrived at the conclusion that Lectures which really teach will never be popular; lectures which are popular will never teach. This Faraday effect is easy to explain: according to N. Bohr, “clearness and truth are in a quantum complementarity relation.” [Tribute to Vladimir Arnold: Arnold in his own words, Notices AMS
given (twice so far), about a half of the participants got A who could understand the outline, and the majority of the rest got B who managed to understand the core. This semester I wish I can be more generous.

**Self-study guide**

(1) The best way to study is not to work when you do not wish to. If you wish to, concentrate on the study at least 15 minutes.³

(2) When you work, work as actively as possible, because effective learning is always ‘active learning’;⁴ Think what you would do if you encounter the problem as the first person in the world.

Since I learned mathematics and physics without attending any course (beyond 200 in the US level) (because I was a wet chemist), I certainly wished to have books with filled details and with all the problems solved. Thus, these lecture notes may be followed without pencil and paper. However, I learned some will power was needed to use such books effectively, because ‘muscle building’ always requires some load. Therefore, always try to guess the next line or step in the derivation/transformation of formulas before reading the lines. Footnotes with * are devoted to the derivation of marked formulas or to more detailed explanations. The reader can regard them as solutions to technical quizzes.

Even with homeworks, there is not enough space to give all the representative elementary problems. Therefore, to augment the book, I urge the reader to **consult the following two problem books:**


I learned thermal physics from these books. All the problems are fully solved, but many of them are not very easy. Try to solve at least the problems in [A] of these books. These books will be (collectively) quoted as Kubo’s problem book (because the original Japanese version is a single book).

59, 378 (2012) The quotation is from p379.
³Because in your brain cells new coding and noncoding RNAs require at least about this order of time to be transcribed.
## Contents

1. Outlook of the course ......................................................... 4
2. Atomic picture of gases ..................................................... 16
3. Introduction to Probability .................................................. 26
4. Law of large numbers ....................................................... 43
5. Maxwell’s distribution ....................................................... 49
6. Mean free path and transport phenomena ............................... 75
7. Brownian motion ............................................................ 91
8. Macrosystems .............................................................. 115
9. Thermodynamics: Principles ............................................... 121
10. Thermodynamics: General consequences .............................. 141
11. More examples of $\Delta S$ .................................................. 161
12. Introduction to statistical mechanics .................................... 177
13. Statistical mechanics of isothermal systems .......................... 190
14. Classical ideal gas and quantum-classical correspondence ............ 210
15. Information and entropy ................................................... 225
16. Specific heat of solid ...................................................... 235
17. How to manipulate partial derivatives .................................... 244
18. Stability, fluctuation, and response ..................................... 258
19. Thermodynamic approach to fluctuations .............................. 272
20. Chemical potential ......................................................... 281
21. Grand canonical ensemble and ideal quantum systems ............... 299
22. Ideal quantum gases at very low temperatures ....................... 318
23. Photons, Phonons and Internal Motions ................................ 336
24. Phases and phase transitions ............................................. 345
25. Spatial dimensionality and interaction range are crucial .......... 358
26. Why critical phenomena are difficult; mean-field theory .......... 371
27. Improving mean field and transfer matrix ............................ 386
28. Symmetry breaking ....................................................... 398
29. First order phase transition .............................................. 405
1 Outlook of the course

Summary
* Science is an empirical endeavor.
* Science and religion have fundamental difference.
* Our world allows microscopic, mesoscopic and macroscopic descriptions.
* The law of large numbers and deviations from the law allow us to understand macroscopic and mesoscopic worlds.

Key words
Three levels of description: Microscopic, mesoscopic, macroscopic

Now, everybody knows that the materials we see around us are made of atoms and molecules. We could even see them by, for example, atomic force microscopes. However, only 50 years ago no one could see atoms. About 100 years ago the existence of atoms was still disputed.

1.1 Atomisms, ancient and modern
The idea that the world is made of indivisible and unchanging minute particles (atomism) is, however, not a very creative idea. After all, it seems that there are only two choices: (i) the world is infinitely divisible and continuous or (ii) the world is made of indivisible units separated by void (and various easy ideas in between).

Some ancient Greek and Indian philosophers reached atomism. Some philosophers may have favored atomism, because it avoided paradoxes associated with continuum

---

5 You must be able to explain these words (hopefully to your lay friends).
6 Are you really sure you can see them today? The answer is not so straightforward, even if it is affirmative.
7 atom ← atomos: a = “not”, tomos = “cutting”
8 Appreciate asking questions} As we will see soon, the ancient atomism is not quite correct as a scientific idea, since important ingredients to make it as a part of natural science are missing due to the limitation of mere philosophical considerations. However, we should appreciate these philosophers for asking the questions that led them to these ideas. We must appreciate those who have asked new questions. In this sense, according to Carlo Rovelli, Anaximander was the first scientist: “I do not wish to overstate the importance of Anaximander. In the end, we know very little about him. But twenty-six centuries ago, on the Ionian coast, somebody opened a new path to knowledge and a new route for humanity.” (C. Rovelli, The first scientist: Anaximander and his legacy (Westholme, Yardley 2007; English version 2011 (translated by M. L. Rosenberg)), location 187, Introduction.
Leucippus (5th c. BCE) is usually credited with inventing atomism in Greece. His student Democritus systematized his teacher’s theory. The early atomists tried to account for the formation of the natural world by means of atoms and void alone. The void space is described simply as nothing, or the negation of body. Atoms are, by their nature, intrinsically unchanging, but can differ in size, shape, position (orientation), etc. They move in the void and can temporarily make clusters according to their shapes and surface structures. The changes in the world of macroscopic objects were understood to be caused by rearrangements of the atomic clusters.

Thus, atomism explains changes in the macroscopic world without creating new substance. Also all the macroscopic phenomena are naturally ephemeral (‘the second law of thermodynamics’). The most decisive difference between the modern atomism and the ancient atomism is that the latter is devoid of dynamics. Indeed, the ancient atomism allowed motions to displace atoms and to change their aggregate states, but no special meaning was attached to movements themselves (quite contrary to the modern thermal motion which we will learn soon).

1.2 Two enemies of empiricism
As noted in 1.1 the modern science has two pillars: the fact-seeking empirical part (in the narrow sense) and the fact-organizing part (based on the phylogenetic learn-

---


10No ‘interatomic’ forces were conceived. That is, it seems that they imagined interactions between contacting bodies (atoms) but they never thought about forces through the void space. Interactions without contact (through void) seem to be a Newtonian novelty as we will see in Lecture 2.

11Since atomism understands that the world orders emerge from rearrangements of atoms, logically this implies that we human beings as natural phenomena are also understood as special arrangements of atoms. Consequently, ancient atomists were critical against institutionalized religions; atomism and secular humanism are rather harmonious as can be seen in Epicurus. If you read Epicurus (e.g., http://www.epicurus.net/index.html), you will realize how ‘modern’ his various views are. S. Weinberg said: “Religion is an insult to human dignity. With or without it, you’d have good people doing good things and evil people doing evil things. But for good people to do evil things, it takes religion.”

12Recall that even the Archimedean mechanics was essentially statics.

13Epicurus grants atoms an innate tendency to move downward through the infinite cosmos. The downward direction is simply the original direction of atomic fall. Interestingly, however, he allows atoms occasionally to exhibit a slight, otherwise uncaused (stochastic!) swerve from their downward path to avoid ‘ordered parallel motion.’
ing). These pillars are vulnerable, if we are not vigilant enough (to check (i) and (ii) in 1.4), to naive versions of ‘just-so empiricism’ and ‘metaphysical influences’. ‘Just-so empiricism’ means just what we observe is a reliable empirical fact. We always need to reflect on what we (can) actually observe (because often what we see is influenced by our metaphysical framework). The metaphysical biases come from Zeitgeist and various traditional ideas including most religious activities. In a certain sense the former may be a more serious threat, because most scientists are unaware of the prejudices they are raised with.\footnote{\textit{Question the Zeitgeist}}

\section*{1.3 What was beyond philosophers’ grasp?}

The idea that everything is made of irreducible units is, as we have just argued, rather natural; if not infinitely divisible, there must be a unit. However, it is hard to identify what the actual unit is without empirical information. Notice that no one ever imagined that we are made of cells\footnote{You should know that the discovery of nucleus by Brown (of Brownian motion) was a key to the proposal of the cell theory by Schwann and Schleiden in 1839.}. Recognize that the \textit{cell theory} is one of the two pillars of biology (the other is Darwinism). We should clearly recognize that this indicates the limitation of philosophers who are not empirical enough. The lack of the idea of ‘molecule’ from the ancient atomism is also an example of this limitation. Perhaps, it is a sign of progress to recognize that the world does not have the structures we ‘naturally’ expect.\footnote{Kepler’s discovery that the circular orbit is not natural may be an example; this was never accepted by Galileo.}

Mechanics is also beyond philosophers’ grasp. Therefore, modern atomism was beyond the reach of any philosopher.

We must respect empirical facts. Science is an empirical endeavor. At the same time, however, as you recognize from the works of Newton, Maxwell, Darwin, and others, ‘pure empiricism’ is not at all enough to do good science.\footnote{That is, as Confucius said: ‘he who learns but does not think, is lost.’ ‘he who thinks but does not learn is in great danger.’ \textit{(Analect, Book 2)}. You can read more excerpts \textit{here}.}

\section*{1.4 What is the fundamental difference between science and religion?}

\begin{itemize}
\item[\footnote{\textit{Question the Zeitgeist}}] We usually believe that smaller scales the more fundamental phenomena. Thus, the study of extremely small scales of the world is regarded as the fundamental physics. How is this really scientific? So far no one (except perhaps myself) has suspected the conflict between (i) and the so-called fundamental physics. Recently, Sabine Hossenfelder eloquently questioned the practice with respect to (ii). See S. Hossenfelder, \textit{Lost in Math: how beauty leads physics astray} (Basic Books, New York, 2018).
\item[\footnote{You should know that the discovery of nucleus by Brown (of Brownian motion) was a key to the proposal of the cell theory by Schwann and Schleiden in 1839.}]
\item[\footnote{Kepler’s discovery that the circular orbit is not natural may be an example; this was never accepted by Galileo.}]
\item[\footnote{That is, as Confucius said: ‘he who learns but does not think, is lost.’ ‘he who thinks but does not learn is in great danger.’ \textit{(Analect, Book 2)}. You can read more excerpts \textit{here}.}]
\end{itemize}
The question what science is does not have any definitive answer. However, its spirit, especially its empirical backbone, consists of
(i) the humility that constantly makes us reflect on whether we really know and whether our methodology and logic are sound, and
(ii) the resultant skepticism.\(^\text{18}\)

This science spirit must be universal beyond us human-beings wherever there are intelligence and conscience; Science is a conscientious and intelligent way of life.

Thus, there is a fundamental conflict between science and religion; the latter demands the unconditional acceptance of certain propositions.\(^\text{19}\) Thus, the faithful can never emancipate himself from the burden of self-deception.

\(^{18}\) Also, the skepticism applied to itself is crucial: to cut the chain of skepticism off at appropriate positions and ‘to experiment.’ I took these statements from Y Oono, *The Nonlinear World* (Springer 2011), mainly Chapter 5.

Unscientific attitudes and political radicalism are correlated: Read M. Rollwage, R. J. Dolan, S. M. Fleming, Metacognitive Failure as a Feature of Those Holding Radical Beliefs, Current Biology 24, 4014 (2018). Radical participants—on both ends of the political spectrum—showed reduced insight into the correctness of their choices.

\(^{19}\) You might say you read even a very recent commentary by Tom McLeish in Physics Today 71(2) 10 (2018) entitled, ‘Thinking differently about science and religion,’ claiming “Maintaining the view that science and religion are in conflict does no one any favors and is hurting science.” Tom is right, IF you do not care about fundamental consistency and integrity of ones intellectual life. The fundamental distinction I pointed out does not show up unless you conscientiously reflect on what you ultimately found your intellectual activity. “Faith is an evil precisely because it requires no justification and brooks no argument.” (R. Dawkins, *The God Delusion* (Houghton Mifflin Company, 2006) Chapter 8).

McLeish points out “Newton himself is testimony to the deep formative role of Christian theology in the rise of experimental and mathematical sciences.” Thus, there is deep and constructive mutual engagement of science and religious belief: “The truth is that throughout most of history, scientific investigation has gone hand in hand with a commitment to theism, at least in the three Abrahamic faiths. It is, sadly, possible to invent conflict where none needs to be.” However, this simply demonstrates that any (wrong) motivation would do (recall Columbus or Magellan); Stefan Zweig wrote in a biography of Magellan, noting that “he planned and acted in honest error”: Let us not underrate the importance of error. Through the promptings of genius, guided by luck, the most preposterous error may lead to the most fruitful of truths. In every branch of science, hundreds of highly important discoveries have been the outcome of erroneous hypotheses. We simply note that Newton was unable to pursue intellectual self-consistency because of the shackles of the Zeitgeist. As you can read in R. Iliffe, *Priest of Nature*, the religious worlds of Isaac Newton (Oxford, 2017) Newton was dead serious about showing that the central Christian doctrine of the Trinity was a diabolical fraud. His atomism is deeply related to this.

It is sad to find such an article approved to be printed in *Physics Today*. 
1.5 How numerous are atoms and molecules?
How many water molecules are there in a tablespoonful (15 cm³) of water? Although we should discuss how to determine the size or mass of an atom (see Section 7), let us preempt the result.

Suppose one person removes one molecule of water at a time from the tablespoonful of water, and the other person uses the tablespoon to scoop out the ocean water to the outer space. If they perform their operations synchronously, starting simultaneously, which person will finish first? With a simple calculation you will realize that the number of molecules in a spoonful of water is comparable (the ratio is less than ca. 3) to the amount of ocean water measured in spoons. Imagine you scoop out water of a 50 m swimming pool. You will not even try to start.

1.6 Why are molecules so small?
Thus, molecules are numerous. They are numerous because they are tiny. Why is an atom so tiny? This is not a meaningful question, however, because being small or being large is only relative; we cannot say whether a 1 m stick is long or short without comparing it with something else. Let us compare our size with the atom size. The above question properly understood is: why is the size ratio between atoms and us so big? Do not forget that we human beings are products of Nature. Therefore, to compare us with atoms does not imply anthropocentric prejudice. Let us try to understand our size relative to atoms.

Large animals, or, more generally, the so-called megaeukaryotes, are often constructed with repetitive units such as segments. The size of the repeating unit is at least one order larger than the cell size. Consequently, the size of ‘advanced’ organisms must be at least 2-3 orders as large as the cell size.

---

20 What if the tablespoons are replaced with teaspoons (5 cm³)?

21 To recognize this trivial fact is the first step to dimensional analysis, an important way of thinking in physics. Read “Introduction to Dimensional Analysis”. In these lectures dimensional analytic explanations will be attempted whenever dimensional analysis can be used.

22 You might recall Protagoras, who said: Man is the measure of all things. However, the original meaning of this statement seems to have been much more restricted, because the word ‘things’ in the original only meant things human beings created (ideas, feelings, social entities, etc., not stars, mountains, etc.). See http://en.wikipedia.org/wiki/Protagoras.
Thus, the problem is the cell size.\textsuperscript{23} We are \textit{complex systems},\textsuperscript{24} so we have our parents and the crucial information and materials required to build us comes from the preceding generation. Since there is no ghost in the world, information must be carried by a certain thing (no ghost principle). Stability of the thing requires that information must be carried by polymers. What polymer should be used? Such a question is a hard question, so we simply imagine something like DNA. ‘No ghost principle’ tells us that organisms require a certain minimal DNA length. This seems to be about 1 m. As a ball its radius is about 0.5 \textasciitilde 1 \mu m. This implies that our cell size (eukaryotic cell size) is \( \sim 10 \mu m \) (= \( 10^{-5} \) m).

Thus, the segment size is about 1mm, and the whole body size is about 1 cm (this is actually about the size of the smallest vertebrates). If we require a good eyesight, the size becomes easily one to two orders more, so intelligent creatures cannot be smaller than \( \sim 1 \) m. That is, the atom size must be \( 10^{-10} \) as large as our size.

We have, at least roughly, understood why atoms are small or why we are big.

1.7 Our world is lawful to the extent of allowing the evolution of intelligence

We have discussed, with the aid of atomism and cell theory, that science is an empirical endeavor and that no correct world view is obtainable solely with philosophical meditations without observing the world. Who observes the world? We observe the world and are making science, so we must be at least slightly intelligent. To be intelligent at least we are \( 10^{9-10} \) as large as the atom. But a large size is not enough. The world must have allowed intelligence to evolve; we must have been able to emerge.

If there is no lawfulness at all, or in other words there is no order in the world,\textsuperscript{25} there is almost no paper discussing the cell size seriously, but recently a relevant paper appeared: Marshall WF, Young KD, Swaffer M, Wood E, Nurse P, Kimura A, Frankel J, Wallingford J, Walbot V, Qu X, Roeder AH, What determines cell size? BMC Biology \textbf{10}, 101 (2012). It is an interesting collection of articles discussing relevant topics to cell size, but no relation with the required information is discussed. However, it has been recognized well that the amount of the DNA in a cell (the so-called C-number) is well correlated with the cell size (see for a summary, T. R. Gregory, “Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma,” Biol. Rev. Camb. Philos. Soc. \textbf{76}, 65-101 (2001)). Thus, we may safely claim that the lower bound of the cell size is determined by the amount of DNA.

\textsuperscript{24}The so-called complex systems studies study spontaneous formation of certain (ordered) structures from disorder. Thus, they study only pseudo-complex systems, because spontaneous emergence is a telltale sign of simplicity; they can emerge spontaneously, because they are in essence simple. In contrast, you did not spontaneously emerge, because to make you was not very simple. Pasteur realized the fundamental complex nature of life: life comes only from life, and never emerges spontaneously within a short time. Notice that no books discussing ‘complexity’ really discuss complexity. See, e.g., Chapter 5 of Y. Oono, \textit{The Nonlinear World} (Springer, 2012).

\textsuperscript{25}‘Order’ may be understood as redundancy in the world; knowing one thing can tell us something
then intelligence is useless; calculation is useless. We use our intelligence to guess what happens next from the current knowledge we have. If in a certain world organisms’ guesses using their intelligence are never better than simple random choices (say, following a dice), then intelligence would not evolve; recall that the human brain is the most energy consuming very costly organ. This means that the macroscopic world (the world we observe directly on our space-time scale) must be at least to some extent lawful with some regularity; we believe in the lawfulness of the world to the extent that we are superstitious.

However, if the law or regularity is too simple, then again no intelligence is useful. If the world is dead calm, no intelligence is needed. The world must be just right (the Goldilocks principle or the principle of moderation). The macroscopic world we experience is not violent but not dull.

1.8 Microscopic world is unpredictable

In contrast, we know the world of atoms and molecules (the microscopic world) is a busy and bustling world. They behave quite erratically and unpredictably (despite deterministic nature of mechanics) by at least two reasons, chaos and external disturbances.

Maxwell clearly recognized that molecules behave erratically due to collisions. Perhaps the simplest model to illustrate the point is the \textit{Sinai billiard}. A hard ball (or rather you can imagine an ice hockey puck) is moving on the flat table, which has a circular obstacle on it. The ball hit the obstacle and is bounced back specularly (see Fig. 1.1).

Roughly speaking, a small deviation of the direction of the particle is doubled upon specular reflection at the central circle, so, for example, to predict the direction of the particle after 100 collisions is very hard. Imagine what happens if there are numerous such particles colliding with each other. Thus, predictions would be abso-

---

26Here, ‘better’ means it is more favorable to the reproductive success of the organisms.

27You will not study if your grade is randomly assigned.

28Its weight is 2% of the body weight, but it consumes about 20% of the whole body energy budget.

29Our logical brain is a reflection of the logical nature of the environment we evolved. This must be parallel to the fact that many fishes have hydrodynamically optimal shapes. Water obeys hydrodynamics, not because fishes swim in it!

30Mistaking correlation as causality is an important ingredient of superstition.

31This is the meaning of the statement appearing later that the world is phenomenologically describable.

32It is convenient to remember that $2^{10} \simeq 10^3$, so $2^{100} \simeq 10^{30}$. 

10
Figure 1.1:  Sinai billiard: **Left:** a motivation. Two hard elastic discs (pucks) are running around on the table with a periodic boundary condition (if a disk disappears from one edge, it reappears from the opposite edge with the same velocity), colliding from time to time with each other. This is a toy model of a confined gas. **Right:** If the dynamic of the center of mass (CM) of one disk is observed from the CM of the other disk, the former may be understood as a ballistic motion of a point mass with occasional collisions with the central circular obstacle. This is called the *Sinai billiard*, and is known to be maximally chaotic.

Thus, the microscopic world is full of noise, and everything looks stochastic, even though the intrinsic mechanics is not at all stochastic. Consequently, it is traditional that the microscopic world is handled with *Kinetic Theory* that grafts space-time local collision dynamics (in many cases binary collision dynamics) and the statistical description of one particle properties (e.g., its position and momentum). This line of approach was developed into quite a sophisticated theory by Boltzmann. The theory allows us to understand time-dependent changes of a system, but since it is very hard to discuss simultaneous multiple collisions, it can study only dilute systems; it is almost hopeless to study condensed matter (e.g., liquid) honestly within the framework of kinetic theory, so in this course only a very elementary introduction to kinetic theory will be given.

1.9 Why our macroscopic world is lawful: the law of large numbers

The world on the scale of atoms is full of noise. We know our scale is quite remote from the atomic scale. The time scales are also disparate; the time scale required

---

33 Quantum mechanically, subtle entanglements are easily lost by perturbation, so the system is much more fragile than the classical counterpart.

34 The latest summary of difficulties may be found in Isabelle Gallagher, From Newton to Navier-Stokes, or how to connect fluid mechanics equations from microscopic to macroscopic scales, Bull. Amer. Math. Soc. **56**, 65-85 (2019).
to describe molecular dynamics is $0.1 \text{ fs} = 10^{-16} \text{ s}$, but the shortest time span we can recognize must be longer than $10 \mu\text{s} = 10^{-5} \text{ s}$. Lawfulness must come from suppression of noise. Our size is crucial to suppress noise; even if particles in a small droplet undergo quite erratic motion, if many particles are averaged, the erratic effect would disappear. This statement may be formally expressed as follows.

Let $X_n$ be random variables. Here, $n$ is the suffix to specify the $n$th variable; we consider a collection of numerous ($N$) such variables, and $X_n$ is the $n$th among them. Then,

$$\sum_{n=1}^{N} X_n = Nm + o[N],$$

(1.1)

where $m$ is the average value (= expectation value) of $X_n$. This is the law of large numbers, the most important pillar of probability theory and the key to understanding the macroscopic world. You may imagine outcomes of coin tossing as an example: $X_n = 1$ if the $n$th outcome is a head; otherwise, $X_n = 0$. By throwing a coin $N$ times, we get a 01 sequence of length $N$, say, 0100101101110101. You can guess the sum is roughly $N/2$, where $N$ must be sufficiently large. This is the law of large numbers. We clearly see the importance of being big (relative to atoms).

1.10 We live in a rather gentle world

You might object, however, that being big may not be enough; we know violent phenomena in the macroscopic world like turbulence or perhaps the cores of galaxies. If the variations are too big, perhaps we may not be able to expect the expectation value to settle within a reasonable narrow range. Also even if the expectation value eventually converges, needed $N$ in the law of large numbers should not be too big; if you can recognize the regularity of the world only after averaging the observations during 1000 generations, probably the law of large numbers cannot favor intelligence very much. Thus, as already discussed above, the world in which intelligence can emerge cannot be too violent. We emerge in the world in which the law of large numbers hold rather easily at large scales to allow macroscopic laws (actually the world close to no change from the molecular point of view). We live in the world where space-time scale is not only quite remote from the microscopic world of atoms.

---

35 We will discuss what we wish to mean by ‘random variables’ more carefully later, but here, you have only to understand them as variables that take various values in an unpredictable fashion.

36 $o$: This standard symbol means higher order small quantities. In the limit being discussed, if $X/Y \to 0$, then we write $X = o[Y]$, which is read: compared with $Y$, $X$ is a higher order small quantity in the limit being discussed. This does not mean $X$ and $Y$ themselves are infinitesimal. For example, $N^{0.99}$ is $o[N]$, if $N$ is large (in the $N \to \infty$ limit), because $N^{0.99}/N = N^{-0.01} \to 0$. 

12
and molecules, but also the extent of nonequilibrium is not too large.\textsuperscript{37}

\section*{1.11 Thermodynamics, statistical mechanics and phase transition}

The macroscopic world close to equilibrium\textsuperscript{38} can be described \textit{phenomenologically} by \textit{thermodynamics}. Here, ‘phenomenologically’ implies that what we observe directly can be organized into a single logical system without assuming any entities beyond direct observations. Thermodynamics is distilled from empirical facts observable on our scale, so it is the most reliable theoretical system we have in physics.\textsuperscript{39}

As we will learn in Lecture 12, statistical mechanics obtains the Helmholtz free energy $A$ (which will be explained in detail later; Lecture 9) as

\begin{equation}
A = -k_B T \log Z, \tag{1.2}
\end{equation}

where $k_B$ is the Boltzmann constant, $T$ is the absolute temperature, and $Z$ is the (canonical) \textit{partition function}

\begin{equation}
Z = \sum e^{-H/k_B T}. \tag{1.3}
\end{equation}

Here, $H$ is the system Hamiltonian (the energy function or energy operator in quantum mechanics) and the summation is over all the \textit{microscopic states}. We will discuss thermodynamics and its relation to statistical mechanics in Lecture 11, and then will learn how to use it subsequently.

$e^{-H/k_B T}$ is a smooth function of $T$ ($>0$), so if the number of terms summed in (1.3) is finite, nothing very singular can happen in $A$ as a function of $T$. However, if the system is very big (ideally, infinitely big, in the so-called \textit{thermodynamic limit}), $A$ can lose smoothness as a function of $T$; the sum of infinitely many smooth functions

\textsuperscript{37}We need a stable simple laws for feeble minds to work (recall the intelligence must evolve). Our macroscopic world is so lawful that some of us can even believe in the benevolence of God.

\textsuperscript{38}Intuitively, you may consider a system is close to equilibrium if all the rapid changes in it have subsided.

\textsuperscript{39}Needless to say, classical mechanics, electromagnetism, quantum mechanics, etc., are also reliable theoretical systems based on our empirical observations. While thermodynamics is used with conscious recognition of its limitations (applicable only to macroscopic systems in equilibrium), other theories are (were) often regarded valid unconditionally (i.e., without clear recognition of their valid domains). In this sense these theories are less reliable. We must learn a lesson from the history that classical electrodynamics was regarded as the ultimate theory until it was recognized not to work in the microscopic world. Now, it is believed quantum mechanics is correct on all scales, and so is the general theory of relativity. Therefore, the current big issue is to unify these two, but we must admit that empirical facts recede from the foreground.
need not be smooth. Thus, *phase transitions* can occur.

Macroscopic world is not generally in equilibrium, so thermodynamics and statistical mechanics cannot describe many macroscopic phenomena. If the extent of nonequilibrium is not too large, appropriately generalized statistical thermodynamic framework is still useful; these include theory of transport phenomena and fluid dynamics. Although these topics are mostly beyond the scope of these lectures, they are actually closely connected to the phenomena we can observe on the mesoscopic scale in equilibrium as will be outlined at the end of this lecture.

### 1.12 Mesoscopic world

What does the world look like if we observe it on the scale intermediate between the microscopic and the macroscopic scales? In (1.1) the $o[N]$ term becomes not ignorable. That is, *fluctuation* cannot be ignored. This is the world where *Brownian motion* dominates, where unicellular organisms live and where the cells making our bodies function. Intelligence is useless, because fluctuation is still too large and prevents agents from predicting what would happen. The best strategy is to wait patiently for a miracle to happen, and if it happens, to cling to it. Molecular motors just do this, crudely put.

In the mesoscopic world, the average of what we observe is consistent with our macroscopic observation results; *Onsager’s regression hypothesis* asserts this. However, if we observe individual systems, observables fluctuate a lot around the expected macroscopic behaviors. Although we will not have time to go into statistical mechanics of such slow macroscopic changes, we will discuss Brownian motion and will give an informal discussion of transport phenomena.

### 1.13 Law of large numbers and probability

We are interested in statistical mechanics, so no one would doubt the relevance of probability theory. What is probability? We will discuss this later (Lecture 2), but let us proceed intuitively. We take statistics, and we know if the number of samples is increased, then statistical results become more reliable. This is just the law of large numbers we have already encountered. The law of large numbers can be written as

$$P \left( \left| \frac{1}{N} \sum X_i - m \right| > \varepsilon \right) \rightarrow 0 \quad (1.4)$$

as $N \rightarrow \infty$, however small positive $\varepsilon$ we choose, where $P$ denotes *probability* of the *event* in the parentheses. That is, if we obtain an empirical expectation value
\[
\frac{1}{N} \sum X_n \text{ using } N \text{ samples, its deviation exceeding } \varepsilon \text{ from the true average value becomes increasingly unlikely as } N \text{ is increased, however small positive } \varepsilon \text{ we choose. If a system is in equilibrium, this limit describes the world of macroscopic equilibrium governed by thermodynamics.}
\]

**1.14 Large deviation and fluctuation**

Now, we ask what happens between the microscopic and macroscopic scales, so we cannot take \(N\) very large. We should study how the above probability goes to zero as a function of \(N\). This is governed by the *large deviation principle*:

\[
P\left(\frac{1}{N} \sum X_i \sim x\right) \sim e^{-NI(x)}, \tag{1.5}
\]

where \(I\) is called the *large deviation function* (or *rate function*), and may be approximated with a quadratic function when \(x\) is close to the true expectation value \(m\):

\[
I(x) \simeq \frac{1}{2V}(x - m)^2. \tag{1.6}
\]

Here, \(V\) is a positive constant (corresponding to variance) and \(m\) is the expectation value, where \(I(m) = 0\) implies the law of large numbers. (1.6) means that mesoscopic noise is usually Gaussian. That is, with the aid of a Gaussian noise \(w\) whose average is zero and variance \(V\), we can write

\[
\frac{1}{N} \sum X_i = m + w. \tag{1.7}
\]

As we will see later (Lecture 18), \(I\) is related to the decrease of entropy from equilibrium due to fluctuations, and the above relation is useful in understanding fluctuations we can observe spatially locally in a system.
2 Atomic picture of gases

Summary

Gay-Lussac gave key empirical facts: $PV \propto T$, the law of constant temperature (for adiabatic free expansion) and the law of combining volumes.

Bernoulli related temperature and (translational) kinetic energy of molecules, but to make kinetic theory precise, we need probability.

Key words
Law of partial pressure, Joule-Thomson experiment, equipartition of energy.

What you should be able to do

Explain the law of constant temperature.

Explain why the Joule-Thomson experiment could kill Newton’s repulsive force model of gases.

Derive the equipartition of energy (for the translational motion).

The following book is recommended for a historical background:

2.1 Aristotelian physics and Galileo’s struggle

According to Aristotle’s (384-322 BCE) physics, the four properties, hot, cold, dry and wet were irreducible properties, which corresponded to four elements of Empedocles (ca 490-430 BCE), fire, water, earth and air, respectively. The crucial point is that what we observe directly by our sense has a direct materials basis.

This type of ideas is called ‘thingification’ or ‘reification.’ Chemistry is naturally under its spell; one might say genomic biology is struggling to emancipate itself

---

40Historical comments in these lectures are heavily dependent on Y. Yamamoto, Historical Development of Thoughts of Heat Theory [in Japanese] (2007-8).
41This summarizes what you should be able to do in practice. Most things required in this course are practical.
42Originally, ‘physica’ meant study of nature.
43Needless to say, the modern chemistry never thinks color and odor are the properties of the substances themselves, but still it tends to explain properties of substances in terms of the properties
from this.\textsuperscript{44}

Even Galileo (1564-1642) was initially under this influence, but later he clearly established the mechanical view of Nature, asserting that what we could feel (e.g., color, odor, etc.) existed only in the relation between the sensing subjects and the sensed objects and was thus subjective and secondary; only the (geometrical) shapes, numbers, configurations (positions) and movements (position changes) of substances were objective and were primary properties.

Mechanics was a key factor for Galileo\textsuperscript{45} to overcome this habit of thought called reification.

Then, you might think Galileo could have invented a kinetic theory of gases and could have conceived warmth as ‘thermal motion.’ This is ‘partially’ correct.

Galileo conceived a special substance ‘fire particles,’ whose vigorous motion was regarded as heat/warmth. It seems that he wished to distinguish ‘microscopic motion’ from ‘macroscopic motion,’ because he did not know the heat engine. The relation between motion and heat was in a certain sense recognized thanks to the fire arms, but it may not be surprising that the relations of heat to the ordinary ‘slow’ motions and to the motions of bullets may not have been identified.\textsuperscript{46}

\subsection*{2.2 Boyle: the true pioneer of kinetic theory of heat}

Boyle (1627-1691) was the first to accept the principle that matter and motion were the primary things, and was truly free from the Aristotelian ‘reificationism.’ He correctly asserted that there were microscopic and macroscopic motions. The former was sensed as heat but could not be sensed by us as motion; the only motion we could sense as such was the ‘progressive motion of the whole’ (i.e., the systematic motion), which could not be felt by us as warmth even if it was vigorous. Thus, Boyle paved the way to the discussion of mutual convertibility of heat and (macroscopic) motion. Boyle was the true pioneer of the motion or the kinetic theory of heat.

\begin{itemize}
  \item of the atoms and molecules more or less directly. For example, if you find an acidic organic compound, you would likely think of COOH.
  \item e.g., such a superstition that there is a gene (as \textit{FoxP2} governing the capability to speak).
  \item Mechanics, or more precisely, studying Archimedes was the key for Galileo and Descartes to overcome the Aristotelian ‘physics’. Archimedes gave them the conviction that the natural laws could be formulated mathematically; indeed the world is mathematically constructed.
  \item Galileo never seems to have paid attention to the frictional heat, but even if he noted this, it would not have been so trivial to go from there to the idea of converting heat to systematic motion. As we will see later, it is doubtful that even Thomson clearly understood the relation when Clausius established thermodynamics.
\end{itemize}
2.3 Discovery of atmospheric pressure and vacua
The biggest discovery of modern physics about gases was the discovery of atmospheric pressure and vacua by Torricelli (1608-1647), Pascal (1623-1662) and von Guericke (1602-1686). This was a discovery demarcating the medieval and the modern ages, its importance only second to heliocentrism. Do not forget that even Galileo explained the impossibility of sucking water more than 10 m in terms of the competition of gravity and the abhorrence of vacua by air.

Within the Aristotelean system, air and fire were regarded essentially light elements, having the tendency to go away from the earth. Therefore, the idea of mass (or weight) of air could not be born. The discovery of vacua decisively discredited Aristotle.

2.4 Daniel Bernoulli and modern dynamic atomism
Thus, a modern dynamic atomic theory should be possible at any time, and indeed, Daniel Bernoulli’s (1700-1782) gas model\textsuperscript{47} (1738\textsuperscript{48}) was the first fully kinetic model. We will discuss a simplified version (ignoring the size of atoms) in a modern fashion shortly.

However, the success of Newtonian universal gravity almost derailed atomism based on mechanics. Bernoulli’s work was forgotten for 100 years.

2.5 Newton derailed kinetic theory of heat completely
Newton (1642-1727) tried to explain Boyle’s law (i.e., $PV = \text{constant}$, where $P$ is the pressure, and $V$ the volume) in terms of (repulsive) forces acting between particles. The idea of forces among particles was a novel idea actually deviating from the tradition of mechanistic theories. For Newton’s contemporary scientists (and also for himself), introduction of gravitational force that explains the solar system was so impressive that the take-home lesson of the Newton’s success was a program to find forces that explain various phenomena; Newton wrote in author’s preface to \textit{Principia}, “I wish we could derive the rest of the phenomena of nature by the same kind of reasoning from mechanical principles; for I am induced by many reasons to suspect that they may all depend upon certain forces by which the particles of bodies, by some causes hitherto unknown, are either mutually impelled towards each other, and cohere in regular figures, or are repelled and recede from each other; which forces being unknown, philosophers have hitherto attempted the search of nature in vain;

\textsuperscript{47}This was in his book on hydrodynamics.
\textsuperscript{48}J S Bach, \textit{Mass in B minor} (BWV 232; about 110 min) was the same year.
but I hope the principles here laid down will afford some light either to this or some truer method of philosophy."  

Crudely put, somehow, Galileo’s fire particle, Newton’s ether, and the ‘pure elemental fire’ of Boerhaave (1668-1738) were understood as analogues. Newton’s repulsive (springy) molecules were imagined due to the clouds of such particles surrounding the molecules.

In any case, for about 100 years, this Newton’s program stifled the kinetic attempts.

You may have been surprised by this episode, but you will learn in your real life how vulnerable the so-called scientists are to current trend/fashion and authority. This is quite unscientific; it is not surprising that the unfathomable gap between science and religion is not recognized.

### 2.6 Between Bernoulli and Maxwell

Between Daniel Bernoulli (ca 1740) and the birth of the modern kinetic theory (due to Maxwell (1831-1879) ca 1860) were the general acceptance of chemical atomic theory (ca 1810) and the birth of physics in the modern sense. Also during this period crucial empirical facts were accumulated, making kinetic theory almost the sole consistent explanation of gasses.

### 2.7 Dalton

Dalton (1766-1844) asserted the law of partial pressure:  
the total pressure of a gas mixture is simply the sum of the pressures each kind of gas would exert if it were occupying the space by itself. As illustrated in Fig. 2.1, it is very naturally explained from the atomic point of view.

---

49 *Principia*, author’s preface (May 8, 1686).

50 Newton’s philosophical starting point was New Platonism of Cambridge and Alchemy; both presupposed that the world is activated by the active principle. Ether was understood as the protoplasm created by God to ‘entrust’ His own activity. Initially, Newton conceived a pan-etherial cosmology.

51 Dalton arrived at his atomic theory not very inductively as is stressed by Brush on p32; Dalton’s writings are sometimes hard to comprehend due to arbitrary thoughts and their outcomes being nebulously mixed up with real experimental results (Yamamoto *loc. cit.* p194), quite different from well-educated Gay-Lussac.
Gay-Lussac (1778-1850) then established three important laws (ca 1810\footnote{1810: Napoleon married Marie Louise of Austria; Chopin was born; Beethoven Piano Trio \textit{Archduke} (Istomin-Stern-Rose)] Notice that Gay-Lussac was the first generation of professional scientist trained professionally to be a scientist. He was a product of French Revolution.}:

(i) The law of thermal expansion of gases (also called \textit{Charles’ law}; $P \propto T$ if $V$ is constant).

(ii) The \textit{law of constant temperature} under adiabatic expansion: if a gas is suddenly allowed to occupy a much larger space by displacing a piston, there is practically no temperature change. You can simulate this nicely using \url{https://phet.colorado.edu/en/simulation/legacy/gas-properties}.

(iii) The \textit{law of combining volumes}: in gas phase reactions the volumes of reactants and products are related to each other by simple rational ratios implying that ‘particles’ cannot generally be atoms.\footnote{But Dalton rejected this interpretation, saying, Gay-Lussac’s experiments were inaccurate, etc. This clearly indicates that Dalton was a metaphysicist more than a physicist.}

\begin{equation}
P = P_A + P_B
\end{equation}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{law_of_partial_pressure}
\caption{The law of partial pressure due to Dalton}
\end{figure}

\section*{2.9 Avogadro}
In 1811\footnote{1811: New Madrid earthquake, J. Austin published \textit{Sense and Sensibility}.} Avogadro (1776-1856) proposed \textit{Avogadro’s hypothesis}: every gas contains...
the same number of molecules at the same pressure, volume, and temperature. However, the molecular theory was not generally accepted until 1860, when Cannizzaro (1826-1910) advocated Avogadro’s proposal in the Karlsruhe Congress (However, Clausius accepted this by 1850;\textsuperscript{55} actually, Cannizzaro was triggered by Clausius’ kinetic theory paper a year before\textsuperscript{56}).

2.10 Daniel Bernoulli’s kinetic theory

\url{http://falstad.com/gas/} is an excellent site to play with a gas dynamic model in java with a heater/cooler, with or without gravity, etc.

Let us look at Daniel Bernoulli’s work. The (kinetic interpretation of) pressure $P$ on the wall is the average momentum given to the wall per unit time and area by the gas. Consider the wall perpendicular to the $x$-axis (see Fig. 2.3).

![Bernoulli's theory](image)

Figure 2.3: Bernoulli’s theory (or mechanical model of gas). Particles are so small that they are assumed not to collide with each other.

Let us proceed step by step. Assume that the mass of each particle is $m$, and that the number density of the particles is $n = N/V$, where $V$ is the volume of the (uniform) gas and $N$ the total number of particles:

(i) For a single particle with velocity $\mathbf{v} = (v_x, v_y, v_z)$ hitting the wall ($v_x > 0$) in the figure, the momentum given to the wall upon collision must be $2mv_x$.

(ii) The total momentum given to the wall in one second is the force on the wall in the $x$-direction, whose magnitude is equal to $PA$, where $A$ is the area of the wall. For a particle moving toward the wall to hit it within the next one second, it must be within distance $v_x$ from the wall. Therefore, to contribute to the pressure the

\textsuperscript{55}Brush p51

particles with the $x$-component velocity around $v_x$ is $> 0$ must be in the volume of $A \times v_x$.

(iii) Let $n(v_x)$ be the number density of the particles with its $x$-component velocity around $v_x$. Then, the contribution of such particles to the pressure (times the wall area) must be $n(v_x) \times A v_x \times 2 m v_x$ according to (i) and (ii).

(iv) Therefore, summing over all the incoming particles, we get

$$P A = \sum_{v_x > 0} 2n(v_x) A m v_x^2.$$  \hfill (2.1)

That is,

$$P = \sum_{v_x > 0} 2n(v_x) m v_x^2 = \sum_{v_x > 0} \frac{2n(v_x) m v_x^2}{n(v_x)} \sum_{v_x > 0} n(v_x) = 2n_+ m \langle v_x^2 \rangle_+,$$  \hfill (2.2)

where $n_+$ is the number of particles with positive $v_x$, and $\langle \rangle_+$ means the average over molecules with positive $v_x$ (to hit the wall).

(v) We do not expect the mean square velocity of the left-going and right-going particles are different, so $\langle v_x^2 \rangle_+ = \langle v_x^2 \rangle$ (henceforth $\langle \rangle$ generally implies averaging, or calculation of expectation values) and $n_+ = n/2$ (just half of the particles move to the right; notice that we have used the law of large numbers!). Therefore,

$$P = n m \langle v_x^2 \rangle.$$  \hfill (2.3)

(vi) Using the isotropy of the gas, we expect $\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle$, so $\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle = 3 \langle v_x^2 \rangle$. Therefore,

$$P = \frac{1}{3} n m \langle v^2 \rangle.$$  \hfill (2.4)

Or, recalling $n = N/V$, we have

$$PV = \frac{2}{3} N \langle K \rangle,$$  \hfill (2.5)

where $K$ is the kinetic energy of a single gas particle. This equation is called Bernoulli’s equation.

Comparing this with the equation of state of an ideal gas $PV = N k_B T$,\(^{58}\)

$$\langle K \rangle = \frac{3}{2} k_B T.$$  \hfill (2.6)

---

\(^{57}\)Precisely we must write it to be in $[v_x, v_x + dv_x]$, but let us proceed as informally as possible.

\(^{58}\)Here, the modern notations are used; what they knew at that time was that $PV \propto NT$, but they could not find $N$. 
2.11 Equipartition of kinetic energy

Let us see that all the particles in a gas consisting of particles with different masses have, on the average, identical translational kinetic energies. That is, (2.6) holds for any particle in a gas mixture (if it is ‘in equilibrium’). As we will learn later, this is almost self-evident, if we know the basic statistical mechanics, but we should also be able to have elementary understanding. It may be inconvenient if you cannot drive a good car at a high speed on a highway, but if you cannot walk, you will not be able to explore the places where nobody has ever been.

Consider a two particle collision process. In equilibrium (i.e., if, on the average, you cannot discern any change),

\[ \langle w \cdot V \rangle = 0, \tag{2.7} \]

where \( w \) is the relative velocity and \( V \) is the center of mass velocity. If we write these in terms of the velocities of two particles \( v_1 \) and \( v_2 \) and their respective masses \( m_1 \) and \( m_2 \), we have

\[ w \cdot V = (v_1 - v_2) \cdot \frac{(m_1 v_1 + m_2 v_2)}{m_1 + m_2} = \frac{(m_1 v_1^2 - m_2 v_2^2) + (m_2 - m_1) v_1 \cdot v_2}{m_1 + m_2}. \tag{2.8} \]

We know \( \langle v_1 \cdot v_2 \rangle = 0 \), so we get the equality of the average kinetic energies.


Notice that Bernoulli’s formula and equipartition of translational kinetic energy imply that even if all the particles in an ideal gas (non-interacting particle system) are with different masses, still the ideal gas law holds.

**Question.** We have demonstrated the equipartition law, but we can give any initial condition to the gas. Do you believe that the equipartition law eventually holds even if the initial condition does not satisfy the law? □
Q2-1. [elementary question]
(1) Let $w$ be the relative velocity of two molecules in a gas (in equilibrium) consisting of pure substance, and $v$ the velocity of a molecule. Show the relation between the root mean squares of $v$ and $w$.
(2) Let us consider an ideal gas mixture. Let $w$ be the relative velocity of the molecule of mass $m$ and the molecule of mass $3m$ in a gas mixture (in equilibrium), and $v$ the velocity of the molecule with mass $m$. Give the relation between the root mean squares of $v$ and $w$.
[Hint: look at our second demo of the equipartition law; one line calculation should be enough.].

Soln.
(1) This we did in the lecture. Let the relative velocity be $w = v_1 - v_2$.

$$\langle w^2 \rangle = \langle (v_1 - v_2)^2 \rangle = \langle v_1^2 + v_2^2 - 2v_1 \cdot v_2 \rangle.$$  

Two different particles are statistically unrelated (we say, as we learn in the next lecture, statistically independent), so $\langle v_1 \cdot v_2 \rangle = \langle v_1 \rangle \cdot \langle v_2 \rangle = 0$, because the gas does not move as a whole. Thus,

$$\langle w^2 \rangle = \langle v_1^2 \rangle + \langle v_2^2 \rangle = 2\langle v^2 \rangle.$$  

That is, the ratio of the root-mean-square velocities is $\sqrt{2}$.

(2) Let $v'$ be the velocity of the particle with mass $3m$. The relative velocity is $w = v - v'$, so with the same reasoning as (1) we have

$$\langle w^2 \rangle = \langle v^2 \rangle + \langle v'^2 \rangle.$$  

Due to the equipartition of kinetic energy

$$\frac{1}{2}m\langle v^2 \rangle = \frac{1}{2}3m\langle v'^2 \rangle \Rightarrow \langle v^2 \rangle = 3\langle v'^2 \rangle.$$  

Therefore,

$$\langle w^2 \rangle = \langle v^2 \rangle + \frac{1}{3}\langle v'^2 \rangle = \frac{4}{3}\langle v'^2 \rangle.$$  

That is,

$$\sqrt{\langle w^2 \rangle} = \frac{2}{\sqrt{3}}\sqrt{\langle v'^2 \rangle}.$$  

Obviously, if one party becomes very heavy, $w$ and $v$ become similar.

Q2-2 [Phys 101 level question]
On a planet for the hydrogen molecule to escape from its surface (to infinity) it requires the surface temperature 320 K. What is the temperature required for methane to escape from the planet surface?

**Soln.**
To escape the planet surface means to get the gravitational potential energy of

\[ U = G m M / R, \]  

(2.9)

where \( M \) is the planet mass, \( R \) the planet radius, and \( m \) the mass of the hydrogen molecule. This must be provided by the kinetic energy \( 3k_B T_{H_2}/2 \). For methane we need \( U = G m' M / R \), where \( m' \) is the methane mass, so \( 3k_B T_{H_2}/2 \times (m' / m) \) is the required kinetic energy. That is \( T = T_{H_2} m' / m = 8 T_{H_2} = 2560 \) K.
3 Introduction to Probability

Summary
* Probability is essentially the ‘volume’ of our confidence measured on the 0-1 scale. Probability must satisfy additivity. Gamble called survival race forces our subjective probability to agree with empirical probability.
* Understand how to describe events in terms of sets.

Key words
Probability, elementary event, sample space, event, conditional probability, (statistical) independent event, stochastic (random) variable, expectation value, variance, standard deviation, indicator, statistical independence of random variables

What you should be able to do
* Be able to calculate expectation values and variances for simple cases.
* Understand $P(A) = \langle \chi_A \rangle$.
* There is an appendix on the elementary combinatorics at the end of the lecture. Be familiar with its content (esp., the binomial theorem and the multinomial theorem).

To go beyond Bernoulli and these elementary discussions, we need the idea of probability. When Maxwell was 19 years old, he read an article introducing the continental statistical theory into British science (e.g., Gauss’s theory), and was really fascinated by it. He wrote to his friend: “the true logic for this world is the Calculus of Probabilities ···.”

Following is an introduction to measure-theoretical probability theory, although no formal introduction of measures will be discussed. You can simply understand that a ‘measure’ is a precise concept corresponding to volumes and weights.

3.1 Probability is a measure of confidence level

---

59 This summarizes what you should be able to do in practice. Most things required in this course are practical.
60 Brush p59
61 An introductory exposition of measure may be found in YO, The Nonlinear World (Springer, 2012) p66.
Suppose we have a jar containing 5 white balls and 2 black balls. What is the degree \( C_w \) on the 0-1 scale of your confidence for you to pick a white ball out of the jar? We expect that, on the average, 5 times out of 7 we will take a white ball out. Hence, it is sensible to say that our confidence in the above statement is \( 5/7 \) on the 0-1 scale; \( C_w = 5/7 \).

![Figure 3.1: Take out one ball without looking in the jar with replacement. ‘How much’ are you sure you get a white ball?](image)

Suppose you can obtain a dollar if you pick a white ball out, but otherwise must pay \( X \) dollars. Whether you wish to participate in this gamble or not depends on \( X \). What is the wise choice? With our confidence \( C_w \) let us assume that we may estimate the expected gain as:

\[
E_M = C_w - (1 - C_w)X. \tag{3.1}
\]

If \( E_M \) is non-negative, i.e., if \( X \leq C_w/(1 - C_w) \), then we may play this gamble. Since you are free to have any idea or thought, you may freely assume \( C_w \) to be any number between 0 and 1. However, there is no freedom of action, if you wish to stay happily in this world. \( C_w \) must be realistic. For our jar game \( C_w = 5/7 \) is demanded. We will soon learn why; we can check whether your confidence level is rational or not empirically by repeating the gamble.

However, probability seems to show up even in cases where we cannot repeat events. How can we check our confidence level is rational or not? For example, the meaning of the statement that the precipitation probability tomorrow is 70% is that we should have a confidence level of 0.7 in raining tomorrow (if you bet money on weather, you’d better use this confidence estimate). However, we cannot repeat ‘tomorrow,’ so how can we check the choice is good? In practice, the extent of confidence is estimated relying on the past experiences of similar events.\(^{62}\)

---

\(^{62}\)Why does the estimated confidence level often match reality? It is thanks to the totality of our 4 billion year experiences (this is called phylogenetic learning). Even if an event does not seem to be repeatable, sufficiently many very similar events happened in the past. “What has been is what will be, and what has been done is what will be done; there is nothing new under the sun.” (Eccles. 1-9) We have been selected to be able to use the result of phylogenetic learning.
3.2 Events and sets
To make mathematical theory of probability, we must express events (= what can happen) as sets. An event which cannot be analyzed further, or need not be analyzed further for our purpose is called an elementary event. Elementary events need not be atomic events that cannot be dissected further into more basic events. For example, when we cast a dice, usually we regard a particular face, 1, 2, 3, 4, 5, or 6 to be up as an elementary events. However, if we pay only attention to the even-odd properties of the numbers, the elementary event could be even and odd. On the other hand, if you wish to use the direction of the edges or locations of the dice as well as the faces, then 1, · · · , 6 are no more elementary events. In statistical mechanics, elementary events are mechanical events that are not dissected further in mechanics.

Denote by Ω the totality of elementary events (called the sample space) allowed in the situation or to the system under study. Any (compound) event under consideration can be identified with a subset of Ω (Fig. 3.2).

When we say an event corresponding to a subset 𝐴 of Ω occurs, we mean that one of the elements (= elementary events) in 𝐴 actually occurs.

![Figure 3.2](image)

Figure 3.2: In this illustration the sample space is Ω = {𝑎, 𝑏, 𝑐, · · · , 𝑥, 𝑦, 𝑧}, where letters denote elementary events; one of the elementary events is what actually happens (or what is actually sampled). Event 𝐴 = {𝑎, 𝑏, 𝑐} is said to occur, if 𝑎, 𝑏 or 𝑐 actually happens.

Events are illustrated conveniently with the aid of the Venn diagrams (Fig. 3.3).\(^6^3\)

3.3 Probability
Let us denote the probability of 𝐴 ⊂ Ω by 𝑃(𝐴). Since probability should measure the degree of our confidence on a 0-1 scale, we demand that

\[ 𝑃(Ω) = 1; \]  

\(^6^3\)due to John Venn (1834-1923) around 1860. who was a logician and a proponent of frequentist interpretation of probability.
something must happen. Then, it is also sensible to assume
\[ P(\emptyset) = 0; \] (3.3)
the event that nothing happens never happens, because something surely happens.

Consider two mutually exclusive events, \( A \) and \( B \). That \( A \) happens implies that one of the elementary events in \( A \) actually occurs. Since \( A \) and \( B \) never occur simultaneously, no elementary event in \( A \) should be in \( B \) (and vice versa). Hence, \( A \cap B = \emptyset \). It is sensible to demand
\[ P(A \cup B) = P(A) + P(B), \text{ if } A \cap B = \emptyset. \] (3.4)
This is the additivity of probability.

For example, for a dice the probability (or your confidence) for 1 is 0.15 and probability for 2 or 3 is 0.4 (needless to say, this dice is not fair or you believe it is not fair), the probability to observe faces with values not more than 3 should be 0.15 + 0.4 = 0.55.

We know other quantities for which additivity (3.4) holds; length, area, volume, mass (if discrete, number), etc. If the probability measures the amount of our confidence, it should be something like volume. A function that assigns numbers to sets (or a map from sets to numbers) is called a set function. Roughly speaking, an additive non-negative set function is a measure. Above examples such as area, volume, etc., are mathematically refined as measures. If a measure whose value on the total set is normalized to unity, it is a probability.

Suppose a shape is drawn in a square of area 1 (a unit square) (Fig. 3.4). If we pepper it with points uniformly, then the probability = our confidence level of a point to land on the shape should be proportional to its area.\(^{64}\) Thus, again it is

\(^{64}\)This has a practical consequence. See 4.7.
intuitively plausible that probability and area or volume are closely related.

Figure 3.4: Peppering the unit square evenly with points, we can estimate the area of $A$.

If an event is given, a certain confidence level in the occurrence of this event may be expressed as a certain probability measure. Whether the confidence level is useful/rational or not is not a concern of probability theory. Probability theory is interested in the conclusions we can deduce from the conditions any confidence belief must satisfy.

As an example, let us consider a series of experiments tossing a coin three times. There are eight possible outcomes corresponding to the combination of three H and T, so the sample space is

$$\Omega = \{\text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}\}.$$  \hfill (3.5)

The word “fair” means that all elementary events are equally likely. Or, it means that we may live without any particular penalty even if we have the same confidence levels 1/8 for the occurrence of any elementary event. However, if one firmly believes that the world is created for H to continue, his confidence level in the occurrence of HHH may be 0.5, and HHT or THH may be 0.2, respectively. Even for such a person the totality of probability must be 1, and the probability for H not to continue must be $1 - 0.5 - 2 \times 0.2 = 0.1$.

The event $A$ that at least two H appear is $A = \{\text{HHH, HHT, HTH, THH}\}$. Since all the elementary events are mutually exclusive, $P(A) = 1/2$ for a person who believes that the coin is fair, but it is obviously larger than 0.9 for the person with a peculiar H belief. The difference between these two confidence levels is so large that very quickly we can check (experimentally) which is realistic.

### 3.4 Relation to combinatorics

As can be seen from the example (especially from the calculation of the confidence levels for the one who believe the coin is fair), in many elementary cases, to count the number of cases satisfying a certain condition is the technical core of probability calculation. However, it is just a technical detail, and is not a crucial part of probability theory. Still, we should be able to do practical calculations, so elementary
Combinatorics is outlined in Appendix 3A.

### 3.5 Objectivity of subjective probability

Since probability is introduced as the confidence level, you might have thought that probability is only subjective. Indeed, in the sense that probability theory is indifferent to whether a particular probability (or confidence level) assignment is useful or not to live in this world, probability may be subjective and not objective. Then, such a subjective concept should not be relevant to objective science such as physics. However, our subjective feeling (emotion underlying decisions) has been molded by natural selection during the past 4 billion years, so our subjective probability estimates (confidence levels) are very often consistent with objective probability.\(^{65}\)

Probabilities appearing in physics should be objective. If we say they are objective, there must be a means to measure them. To this end, we must learn elementary probability theory a bit further.

### 3.6 Review of probability concept

We continue an introduction to probability, so let us review elementary definitions and facts.

* An event \(A\) is understood as a subset of the sample space \(\Omega\), the set of totality of all possible elementary events. That event \(A\) happens implies that one of the elementary events \(\in A\) actually happens.

* The probability \(P(A)\) of event \(A\) is a measure of your confidence level in its occurrence. Additivity holds for probability and \(P(\Omega) = 1\).

   From this, for example, we can derive

   \[
P(A \cup B) = P(A) + P(B) - P(A \cap B).
   \]

   (3.6)

Probability theory does not care whether a probability is objective or not. However, as discussed before, it is often possible to distinguish realistic probabilities and wishful probabilities; our subjective feelings have been selected so that our subjective probability and the actual objective probability are consistent by the past 4 giga year selection process.\(^{66}\)

\(^{65}\)Even other animals have considerable capability of estimating probabilities; they are free from strange religious beliefs, so their capability could be better than ours when clouded by strange beliefs.

\(^{66}\)It is well documented that other animals also have considerable probability-theoretical competence. It might even be the case that their capability is better than ours, since they can not have outrageous beliefs.
3.7 ‘Subadditivity’ of probability

\[ P(A \cup B) \leq P(A) + P(B), \]

(3.7)

and

\[ A \subset B \Rightarrow P(A) \leq P(B). \]

(3.8)

Denoting \( \Omega \setminus A \) by \( A^c \) (complement), we get

\[ P(A^c) = 1 - P(A). \]

(3.9)

**Exercise.** Show, if \( P(A) = 1 \), then \( P(A \cap B) = P(B) \). [Obvious! It is important to feel that this is obviously true, but you should be able to give a logical proof as well. Notice that \( P(B) = P(A \cap B) + P(A^c \cap B) \). From (3.8) \( P(A^c \cap B) \leq P(A^c) = 0 \), so \( P(B) = P(A \cap B) \).]

3.8 Conditional probability

Suppose we know for sure that event \( B \) has occurred. Under this condition what is the probability of the occurrence of event \( A \)? Thus we need the concept of conditional probability. We write this conditional probability as \( P(A|B) \), and define it as

\[ P(A|B) = \frac{P(A \cap B)}{P(B)}, \]

(3.10)

so that \( P(B|B) = 1 \) should hold.

3.9 Statistical independence

When the occurrence of event \( A \) does not tell us anything new about event \( B \) and vice versa, we say two events \( A \) and \( B \) are (statistically) independent. Do not confuse ‘independent events’ and ‘mutually exclusive events.’ Since knowing about event \( B \) does not help us to obtain more information about event \( A \) if \( A \) and \( B \) are independent, we should get

\[ P(A|B) = P(A), \]

(3.11)

where \( P(A|B) \) is the conditional probability just introduced in 3.8. Therefore, the following formula must be an appropriate definition of independence of events \( A \) and \( B \):

\[ P(A \cap B) = P(A) \cdot P(B). \]

(3.12)
For example, when we use two fair dice ‘a’ and ‘b’ and ask the probability for ‘a’ to exhibit a number less than or equal to 2 (event \( A \)) and ‘b’ a number larger than 3 (event \( B \)), we have only to know the probability for each event

\[
\begin{align*}
A & = \{1a, 2a\} \\
B & = \{4b, 5b, 6b\}
\end{align*}
\]

where \( n_x \) denotes the elementary event that dice \( x \) gives face \( n \). Thus, the answer is

\[
P(A \cap B) = P(A) \cdot P(B) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}
\]

for fair dice.

### 3.10 Stochastic variables

You must have heard of ‘stochastic processes.’ A stochastic process is a process in which a ‘stochastic variable’ takes various values as a function of time. Then, what is a ‘stochastic variable’?

Let \( \Omega \) be a sample space and a probability \( P \) is given on it.\(^{67} \) Then, a function (map) from \( \Omega \) to some mathematical entity (real numbers, vectors, etc.) is called a stochastic variable or random variable.

Let \( \Omega = \{ \omega_i \} \). A real-valued stochastic variable \( F \) is a map \( F : \Omega \to \mathbb{R} \). It is rational to write the probability for this stochastic variable to take a particular value \( f \) as

\[
\text{Prob}(F = f) = P(\{ \omega | F(\omega) = f \}) = P(F^{-1}(f)).
\]

(3.13)

Since \( F^{-1}(f) \) is the totality of the elementary events \( \omega \) such that \( F(\omega) = f \), summing all the probabilities for these elementary events should be the probability of the set \( F^{-1}(f) = \text{‘event such that } F = f \text{’} \) (see Fig. 3.5). Therefore, the above definition is very reasonable.

![Figure 3.5: The probability for a stochastic variable F to assume a particular value f](image)

For example, suppose you cast a dice (\( \Omega = \{1, 2, 3, 4, 5, 6\} \)), and you obtain $1 if the face is odd; otherwise, you must pay $1. Then, your gain \( F \) is a random

---

\(^{67}\)Here, that \( P \) is given on \( \Omega \) implies that the value of \( P \) is given for all the elementary events in \( \Omega \) (if \( \Omega \) is discrete; if it is continuous, then \( P \) must be defined on an appropriate family of subsets of \( \Omega \)). \((\Omega, P)\) is called a probability space. If you read a respectable probability book, you will encounter something like \((\Omega, \mathcal{B}, P)\), where \( \mathcal{B} \) is a family of ‘measurable sets.’ We will not discuss this in these notes. (Not all the events should have probabilities to avoid something like \(1 + 1 = 3\), so we must specify what events can have probabilities. This is the role of \( \mathcal{B} \).)
variable \( F : \Omega \rightarrow \{-1, +1\} \) such that \( F^{-1}(-1) = \{2, 4, 6\} \) and \( F^{-1}(+1) = \{1, 3, 5\} \). Therefore, \( \text{Prob}(F = +1) = P(\{1, 3, 5\}) \) and \( \text{Prob}(F = -1) = P(\{2, 4, 6\}) \).

In short, if we write the probability for a stochastic variable \( F \) to assume a value \( f \) as \( P_F(f) \), then
\[
P_F(f) = P(F^{-1}(f)).
\]
(3.14)

### 3.11 Expectation values

The expectation value ( = average) of \( F \) is written as (if one wishes to express the underlying probability \( P \) explicitly) \( E_P(F) \) or \( \langle F \rangle_P \) and is defined by
\[
E_P(F) \equiv \langle F \rangle_P \equiv \sum_{\omega \in \Omega} P(\omega)F(\omega) = \sum_f P_F(f)f.
\]
(3.15)

Often the suffix \( P \) is omitted. The last equality can be checked by a straight-forward calculation (also see Fig. 3.5 above). Let us denote the event \( F = f \) as \( ev(F = f) = \{\omega | F(\omega) = f\} \):

\[
\sum_{\omega \in \Omega} P(\omega)F(\omega) = \sum_f \left( \sum_{\omega \in ev(F = f)} P(\omega) \right) F(\omega) = \sum_f \left( \sum_{\omega \in ev(F = f)} P(\omega) \right) F(\omega) = \sum_f P(ev(F = f))f = \sum_f P_F(f)f.
\]
(3.16)

At the last step the definition of \( P_F \) was used.

The sum becomes integration when we study events which are specified by a continuous parameter. In this case,
\[
E_P(F) \equiv \langle F \rangle_P \equiv \int_{\omega \in \Omega} F(\omega)P(d\omega) = \int_{\omega \in \Omega} F(\omega)dP(\omega),
\]
(3.17)

where \( P(d\omega) \) is the probability of the volume element \( d\omega \); often \( P(d\omega) \) is written as \( dP(\omega) \). You may simply interpret this integral just as the Riemann integral.

\( E \) may be understood as an operator.\(^68\) Let \( f \) and \( g \) be stochastic variables, and

---

\(^{68}\) *Operator* is a map that maps a function to another function or number. For example, the differential operator \( d/dx \) maps a differentiable function \( f \) to its derivative \( f' \) and is a linear operator.
\(a\) and \(b\) real numbers. Then we have the following equality
\[
E(af + bg) = aE(f) + bE(g).
\] (3.18)
That is, the expectation value of a linear combination is a linear combination of expectation values. An operator with this property is called a linear operator. The expectation value operator \(E\) is a linear operator.

### 3.12 Variance
We are also interested in the ‘spread’ of the variables. Its good measure is the variance of \(X\) defined as
\[
V(X) = E([X - E(X)]^2) = E(X^2) - E(X)^2.
\] (3.19)
Its square root \(\sigma(X) = \sqrt{V(X)}\) is called the standard deviation of \(X\).

### 3.13 Indicator
The indicator \(\chi_A\) of a set (= event in our context) \(A\) is defined by
\[
\chi_A(\omega) \equiv \begin{cases} 
1 & \text{if } \omega \in A, \\
0 & \text{if } \omega \not\in A.
\end{cases}
\] (3.20)
This indicates the answer ‘yes’ or ‘no’ to the question: is an elementary event \(\omega\) in \(A\)? \(\chi_A = 1\), if \(A\) happens.

Notice that (apply (3.15) straightforwardly)
\[
\langle \chi_A \rangle_P = \sum_\omega \chi_A(\omega)P(\omega) = \sum_{\omega \in A} P(\omega) = P(A).
\] (3.21)
This is a very important relation for the computation of probabilities.

A random variable (= stochastic variable) is a function \(X(\omega) (\omega \in \Omega)\) defined on \(\Omega\). If we denote the event \(X = x\) as \(ev(X = x) = \{\omega \mid X(\omega) = x\}\), then \(X\) defined on \(\Omega\) may be written as
\[
X(\omega) = \sum_x x\chi_{ev(X=x)}(\omega).
\] (3.22)
(3.15) follows from this and (3.21):
\[
\langle X(\omega) \rangle = \sum_x x\langle \chi_{ev(X=x)} \rangle = \sum_x xP_X(x).
\] (3.23)

35
where $P_X(x)$ is the probability for $X$ to be $x$. Here, we have used the fact that the expectation value operator $\langle \cdot \rangle$ is a linear operator and a similar calculation as (3.16):

$$
\langle \chi_{ev(X=x)}(\omega) \rangle = \sum_{\omega \in \Omega} \chi_{ev(X=x)}(\omega) P(\omega) = P(ev(X=x)) = P_X(x) \quad (3.24)
$$

### 3.14 Independence of stochastic variables

How should we define ‘independence’ (statistical independence) of two stochastic variables $X_1$ and $X_2$? A reasonable answer is that

$$
E(F(X_1)G(X_2)) = E(F(X_1))E(G(X_2)) \quad (3.25)
$$

holds for any functions $^69$ $F$ and $G$ of the stochastic variables. In particular, if stochastic variables $X_1$ and $X_2$ are independent,

$$
E(X_1X_2) = E(X_1)E(X_2). \quad (3.26)
$$

If random variables $X$ and $Y$ are independent, then

$$
V(X + Y) = V(X) + V(Y). \quad (3.27)
$$

### 3.15 Covariance

If you have two random variables, you might wish to know their relations. For two stochastic variables $X$ and $Y$

$$
C(X, Y) = E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y) \quad (3.28)
$$

is called the covariance between $X$ and $Y$, which shows up often when we wish to study fluctuations.

If $X$ and $Y$ are statistically independent variables, then $C(X, Y) = 0$, but the converse is not true. Let $Y = \pm X$, where $\pm$ is randomly chosen by coin-tossing. Then, $C(X, Y) = 0$, but we always have $X^2 = Y^2$, so they cannot be statistically independent; they violate the definition of statistical independence, and also intuitively we cannot say $X$ and $Y$ are unrelated.

$^69$: Any functions’ here means ‘any (Lebesgue) integrable functions.’
Appendix 3A: Rudiments of combinatorics

As noted above, often evaluation of elementary probabilities boils down to counting the number of ways to arrange objects. In statistical mechanics we must be able to count the number of elementary events (i.e., microscopic events) under various constraints. How to count is the main topic of combinatorics.

Sequential arrangement (without repetition) of $r$ objects from $n$ distinguishable objects: $nP_r$

Suppose there is a set of $n$ distinguishable objects. How many ways are there to make sequential arrangements of $r$ objects taken from this set (without repetition)? This number is denoted by $nP_r \equiv P(n, r)$.

There are two ways to get an explicit formula for this number:

(i) There are $n$ ways in selecting the first object. To choose the second object, there are $(n-1)$ ways, because we have already taken out the first one. Here, the distinguishability of each object is crucial. In this way we arrive at

$$P(n, r) = n \cdot (n - 1) \cdots (n - r + 1) = \frac{n!}{(n - r)!},$$

where $n! = 1 \cdot 2 \cdot 3 \cdots (n - 1) \cdot n$; $n$ factorial is the number of ways $n$ distinguishable objects can be arranged in a sequence. The following symbol is also often used:

$$(n)_r \equiv n \cdot (n - 1) \cdots (n - r + 1).$$

(ii) The other derivation is an interpretation of the rightmost formula in (3.29). We can imagine distinguishable objects as monomers and try to make a polymer of length $n$ from these monomers. There are total $n!$ different configurations (different polymers). Now, let us classify these polymers according to the first $r$ monomer arrangements. How many different polymers with a given first $r$ subpolymer? There are $(n-r)!$ ways to complete this subpolymer into a full length $n$ polymer. Therefore, if we classify length $n$ polymers according to the initial $r$ monomer configuration, there are $n!/(n-r)!$ kinds.

Also from the logic in (i), we know that the number of ways to arrange $r$ objects taken from $n$ distinguishable objects with repetition allowed is $n^r$. We can show $(n)_r/n^r \to 1$, if $n$ becomes large with fixed $r$. That is, asymptotically the samplings with and without replacement are the same (as intuitively expected).

\footnote{W. Feller, An Introduction to Probability Theory and Its Applications (Wiley, 1957) volume 1, Chapter II is a useful reference.}
Selection of $r$ objects from $n$ distinguishable objects: binomial coefficient, $nC_r$

Under the same distinguishability condition, we now disregard the order in the arrangement of $r$ objects. That is, we wish to answer the question: how many different subsets can we make, if we choose $r$ elements without repetition from a set consisting of $n$ distinguishable elements?

Since we disregard the ordering in each arrangement of $r$ distinguishable objects, the answer should be

$$nC_r \equiv \binom{n}{r} \equiv \frac{nP_r}{r!} = \frac{n!}{(n-r)!r!}. \quad (3.31)$$

The number $\binom{n}{r}$ is called the binomial coefficient due to a reason clear from (3.34) below.

**Exercise 1.** Show the following equalities and give combinatorial explanations:

$$nP_r = \binom{n}{r} \cdot rP_r, \quad (3.32)$$

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}. \quad (3.33)$$

□

**Binomial theorem**

Consider the $n$-th power of $x + y$. There exists an expansion formula called the binomial expansion:

$$(x + y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r. \quad (3.34)$$

This can be seen easily as follows: We wish to expand the product of $n$ $(x + y)$:

$$\underbrace{(x + y)(x + y)(x + y) \cdots (x + y)}_{n}. \quad (3.35)$$

As an example take the term $x^2 y^{n-2}$. To produce this term by expanding the above product, we must choose 2 $x$’s from $n$ $(x + y)$. There are $\binom{n}{2}$ ways to do this, so the coefficient must be $\binom{n}{2}$.

**Multinomial coefficient**

Suppose there are $k$ species of particles. There are $q_i$ particles for the $i$-th species. We assume that the particles of the same species are not distinguishable. The total
number of particles is \( n \equiv \sum_{i=1}^{k} q_i \). How many ways are there to arrange these particles in one dimensional array?

If we assume that all the particles are distinguishable, the answer is \( n! \). However, the particles of the same species cannot be distinguished, so we need not worry which \( i \)-th particle is chosen first. Hence, we have over-counted the number of ways by the factor \( q_i! \) for the \( i \)-th species. The same should hold for all species. Thus we arrive at

\[
\frac{n!}{q_1!q_2! \cdots q_{k-1}!q_k!},
\]

(3.36)

This is called the *multinomial coefficient*.

**Multinomial theorem**

There is a generalization of (3.34) to the case of more than two variables and is called the *multinomial expansion*:

\[
(x_1 + x_2 + x_3 + \cdots + x_m)^n = \sum_{q_1+q_2+\cdots+q_m=n, q_i \geq 0} \frac{n!}{q_1!q_2! \cdots q_m!} x_1^{q_1}x_2^{q_2} \cdots x_m^{q_m},
\]

(3.37)

whose demonstration is very similar to that explained around (3.35).

**Arrangement of indistinguishable objects into distinguishable boxes**

Consider \( n \) indistinguishable objects. We wish to distribute them into \( r \) distinguishable boxes. How many distinguishable arrangements can we make?

Since the boxes are distinguishable, we arrange them in a fixed sequence, and then distribute the indistinguishable objects (Fig. 3.6).

Hence, the problem is equivalent to counting the number of arrangements of \( n \) indistinguishable balls and \( r - 1 \) indistinguishable bars on a line (Fig. 3.6 bottom). Apply
(3.36) to obtain the answer:

\[ \frac{(n + r - 1)!}{n!(r - 1)!} = \binom{n + r - 1}{n}. \]  

(3.38)

How about the arrangement of the distinguishable \( n \) into \( r \) distinguishable boxes? The first particle can be put into one of \( r \) boxes. Then, the second, etc. Thus, there are \( r^n \) ways.

There are two more conceivable cases:
(i) How about the arrangement of \( n \) distinguishable particles into \( r \) indistinguishable boxes? This is easy.
(ii) How about the arrangement of \( n \) indistinguishable particles into \( r \) indistinguishable boxes? [This is not easy. This is related to the decomposition of \( n \) into \( r \) positive integers = integer partition problem. http://en.wikipedia.org/wiki/Partition_(number_theory)]

**Exercise 2.** How many ways are there to distribute \( n \) distinguishable balls into \( n \) distinguishable boxes?

**Exercise 3.** How many ways to distribute \( n \) distinguishable balls into \( n \) distinguishable boxes with exactly one box left empty? \( [n \choose n] \)

**Exercise 4.** There are 4 workers who produced total 4 defective products. What is the probability of a particular person produced 3 defective products? Assume all the workers are equally skilled.

**Derangement**

A **derangement** is a permutation of the elements of a set such that none of the elements appear in their original position [http://en.wikipedia.org/wiki/Derangement].

Let \( D_n \) be the number of derangements of \( n \) (distinguishable) objects. Then,

\[ D_n = (n - 1)D_{n-1} + (n - 1)D_{n-2} \]  

(3.39)

for \( n \geq 3 \). Note that \( D_1 = 0 \), \( D_2 = 1 \). This gives

\[ \frac{1}{n!}D_n = \sum_{k=2}^{n} \frac{(-1)^k}{k!} = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + (-1)^n \frac{1}{n!}. \]  

(3.40)

This converges to \( 1/e \) in the large \( n \) limit.
Q3-1. [Fun problems]
(1) Which probability is larger, (a) or (b), assuming the 6-sided dice are fair?
(a) At least one ‘1’ face appears in one throw of 4 dice.
(b) Two ‘1’ faces appear simultaneously at least once in 25 throws of two dice.

(2) There are two kittens. You are told that at least one of them is a male. What is the probability that the two kittens are both males? What is the probability that one kitten is a female? (Assume that the sex ratio of kittens is 1 to 1.)

(3) There are 5 boxes A-E of which one contains a prize of $1000. You are asked to choose one box. After you choose one of the five boxes, the ‘coordinator’ of the gamble opens 3 of the remaining boxes which are all empty. Then, he tells you that if you pay $250 you may switch your choice. What is a good choice for you (assuming that you wish to get more money), and your expected gain?

SoLn.
(1)
(a) The complement of the event ‘at least one’ is ‘none.’ That is, if we compute the probability $p$ for the event that no ‘1’ face appears in one throw of 4 dice, $1 - p$ must be the answer ($P(A) = 1 - P(A^c)$). This is $1 - (5/6)^4 = 0.5177$.

(b) This is very similar to (a). The probability $p$ of the complement (no simultaneous ‘1’ is $(35/36)^{25}$). Therefore, $1 - (35/36)^{25} = 0.5055$. Thus, (a) is slightly more likely.

The French nobleman and gambler Chevalier de Méré suspected (purely empirically, of course) that (a) was higher than (b) with 24 throws (in this case the probability is 0.4914; about 5% difference) instead of 25, but his mathematical skill was not great enough to demonstrate why this should be so. He posed the question to Pascal, who solved the problem and proved de Méré correct.\footnote{basically from Wikipedia.} We did better: even if we throw 25 times, still (a) is more likely.

(2) For two kittens (you must recognize kittens can be distinguished), there are 4 different sex combinations: mm, mf, fm, ff. You know one of three occurred: mm, mf or fm, since one is male. These three cases occur with equal probability. Therefore, with 1/3 of the probability the other is male.

(3) If you do not switch your choice, obviously your expected gain will be $200 = 1000 \times (1/5)$.

The remaining 4 boxes contain the prize with probability 4/5. After the coordinator opens 3 empty boxes, this probability is ‘concentrated in’ the remaining box. Therefore, if you switch your choice, then your expected gain would be $800. Thus, definitely you should pay $250 and switch!

In this case, the new information changes the condition, under which you should
reconsider your ‘confidence level.’

The above solution assumes (as usual in this type of questions) that the coordinator knows where the prize is. What if the coordinator does not know where the prize is? Or you can imagine a wind opens three boxes, and they happen to be all empty. In this case, what is your choice? An elementary answer is as follows: (1) If you choose the prize-containing boxes, there are $\binom{3}{1}$ ways to open three boxes. (2) If you choose the empty box, then there is only one way to choose three empty boxes from the remaining 4 boxes. There are 4 ways for you to choose an empty box. That is, both are equally probable. Thus, you should not switch the box.

Suppose you are told that your choice does not contain the prize, but then told that if you pay $300, you may choose a new box. Will you pay this price?

It may be fun to read: J. Rosenhouse, The Monty Hall Problem (Oxford, 2009).

**Q3-2.** There were 400 students in an exam. A professor was interested in who were cheating in the exam, so he watched out for rare agreements of wrong solutions. He found a pair whose errors agreed exactly, and he could calculate that the probability of this agreement was at most $10^{-5}$. Therefore, he accused the pair of cheating. Is his decision rational?

**Soln.**
The number of pairs is $\binom{400}{2} = 200 \times 399 = 79800 \simeq 80000$, so on the average we always find at least 0.8 rare-agreement pair. What does this mean? Without any ill intention, on the average almost one pair will be accused! Therefore, his decision is irrational.

When there are numerous (statistically independent) samples simultaneously as in most problems of bioinformatics (say, ca. 20,000 genes for us), we must be very careful about false positives. There are many ways to cope with this problem, but the simplest is the Bonferroni correction. Look this up.
4 Law of large numbers

Summary
* The law of large numbers allows us to measure probability.
* We can use the law of large numbers to estimate integrals. Recognize the power of randomness.

Key words
Chebyshev’s inequality, law of large numbers

What you should be able to do
* You must be able to use the law of large numbers to estimate how many samples you need to determine the empirical average with a prescribed error tolerance level.

4.1 How can we measure probability?
Let us return to the problem in Fig. 3.4. Suppose $\chi_A$ is the indicator of the area $A$ in the unit square. We pepper dots on it evenly. If the $i$th dot (location $x_i$) is on $A$, $\chi_A(x_i) = 1$, otherwise, 0. If we count the number $N_1$ of points for which $\chi_A = 1$ in the total trial with $N$ dots, $N_1/N$ should be close to the area, which is the probability $P$ for the dot to land on $A$. We expect

$$\frac{1}{N} \sum_{i=1}^{N} \chi_A(x_i) \to E(\chi_A) = P(A) \quad (4.1)$$

in the large $N$ limit. This can be verified by the most important theorem of probability theory: the law of large numbers.

Discussion 1. Suppose we have a fair coin. Which is easier to realize, 6 H out of 10 trials or 55 H out of 100 trials? □

For a fair coin let $X_n$ be the indicator of head (i.e., $X_n = 1$ if the outcome of the $n$th tossing is a head, otherwise, 0) for the $n$-th tossing of the coin. Then, we expect

$$\frac{1}{N} \sum_{n=1}^{N} X_n \to \frac{1}{2}. \quad (4.2)$$

Jakob Bernoulli (1654-1705) proved this (Fig. 4.1).\(^72\)

\(^72\)published posthumously in Art Conjectandi (1713). $e$ was introduced by him as well.
Figure 4.1: Left: The law of large numbers illustrated: the percentage of H in N trials [from http://www.mathaholic.com/?tag=law-of-large-numbers. This article, ‘Why casinos don’t lose money,” is recommended.]. Right: Jacob Bernoulli (1654-1705) proved the law of large numbers (published posthumously in 1713) [Swiss stamp in 1994, courtesy of Professor M. Börgens of Technische Hochschule Mittelhessen].

The most important message is that probabilities of events may be observed experimentally. Although we introduced probability $P(A)$ as a measure of our (subjective) confidence in the occurrence of event $A$, whether the probability is realistic or not can be determined empirically in many cases. Do not forget that our intuition/emotional judgement is based on our nervous systems, which have been subjected to rigorous selection processes in the past 1 billion years. Thus, inevitably, our subjective judgements tend to be consistent with the objective world.

The following URL illustrates the law of large numbers:
http://demonstrations.wolfram.com/IllustratingTheLawOfLargeNumbers/
You will realize that convergence is not very fast.

4.2 Precise statement of law of large numbers
A precise statement of the law of large numbers (LLN) is as follows:
Let $\{X_i\}$ be a collection of independently and identically distributed (often abbreviated as iid) stochastic variables. For any $\varepsilon > 0$,

$$\lim_{N \to \infty} P\left( \left| \frac{1}{N} \sum_{n=1}^{N} X_n - E(X_1) \right| > \varepsilon \right) = 0 \quad (4.3)$$

holds under the condition that the distribution of $X_i$ is not too broad: $E(|X_1|) < \infty$. If $V(X_1) < +\infty$, the condition is satisfied.\footnote{Since all $X_n$ are distributed identically, we use $X_1$ as a representative, so $E(|X_1|)$, etc., show up in the statement.} In the following, the law of large numbers is demonstrated under this assumption.

73
The following is also a precise expression:

\[ \sum_{n=1}^{N} X_n = N E(X_1) + o[N]. \] (4.4)

The interpretation of (4.3) is as follows: we perform a series of \( N \) experiments to produce the empirical expectation value \((1/N) \sum_{n=1}^{N} x_n\). This set of \( N \) experiments is understood as a single ‘run,’ and we imagine many such runs. Then, (4.3) tells us that the probability that these runs produce empirical averages \( S_N/N \) deviating from the true mean \( E(X_1) \) by more than (any positive number) \( \varepsilon \) goes to zero in the limit of the infinite run length.

Remark: Suppose you find an empirical average \( S_N/N \) larger than \( E(X_1) \). Then, you might expect more outcomes smaller than \( E(X_1) \) in the near future. This is the famous gambler’s fallacy (or fallacy of the maturity of chances). See http://en.wikipedia.org/wiki/Gambler’s_fallacy, especially, psychology behind the fallacy. □

4.3 Why is the law of large numbers plausible?
Before going to a rigorous demonstration of LLN, let us understand why it is plausible. We could expect that the average of \( S_N/N \) (the empirical average) should fluctuate around \( E(X_1) \). Its width of fluctuation must be evaluated by the variance: (notice that \( V(cX) = c^2 V(X) \) and ‘additivity’ (3.27) in the following calculation)

\[ V \left( \frac{1}{N} \sum_{n=1}^{N} X_n \right) = \frac{1}{N^2} V \left( \sum_{n=1}^{N} X_n \right) = \frac{1}{N^2} \sum_{n=1}^{N} V(X_n) = \frac{1}{N} V(X_1). \] (4.5)

Thus, the width of the distribution shrinks as \( N \) is increased. That is why \( S_N/N \) clusters tightly around \( E(X_1) \) as \( N \to \infty \). This is the essence of LLN. This is illustrated in:
http://demonstrations.wolfram.com/ChebyshevsInequalityAndTheWeakLawOfLargeNumbersForIidTwoVect/

4.4 Chebyshev’s inequality and a proof of LLN

Note for the instructor: (4.3) and this (if properly stated as an almost sure convergence) are different; the former is called the weak law of large numbers and the latter the strong law of large numbers. However, in many realistic situations we encounter in statistical mechanics, whenever the weak law holds, so does the strong law.
The key to an honest proof of LLN is Chebyshev’s inequality\(^75\)

\[ a^2 P(|X - E(X)| \geq a) \leq V(X). \quad (4.6) \]

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).\(^76\) Fig. 4.2 illustrates the demonstration.

![Figure 4.2](image)

Figure 4.2: (4.9) illustrated. If we introduce the density distribution function \(f\), (4.9) reads

\[ V(X) = \int X^2 dP(\omega) \geq \int_{|X| \geq a} X^2 f(X) dX \geq a^2 \int_{|X| \geq a} f(X) dX = a^2 P(|X| \geq a). \quad (4.8) \]

On the integration range \(|X| \geq a\), \(X^2 \geq a^2\), so

\[ \int_{|X| \geq a} X^2 dP(\omega) \geq \int_{|X| \geq a} a^2 dP(\omega) = a^2 \int_{|X| \geq a} dP(\omega). \quad (4.9) \]

This implies

\[ V(X) \geq a^2 P(|X| \geq a). \quad (4.10) \]

Since we have shifted \(X\) by \(E(X)\), this implies Chebyshev’s inequality (4.6).

\(^75\)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.

\(^76\)Or, you can use \(V(X) = V(X - a)\) for any number \(a\); the width does not change wherever the distribution is placed.
4.5 Proof of the law of large numbers
We wish to apply Chebyshev’s inequality (4.6) to the sample average $(1/N) \sum X_n$. Replacing corresponding quantities in (4.6) $(X \to (1/N) \sum X_n, a \to \varepsilon)$, and using (4.5), we get

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

(4.11)

Taking $N \to \infty$, we arrive at LLN.

4.6 Detecting unfair coins
We have shown that indeed (4.1) can be used to observe the probability of an event. 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 (4.11) implies

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

(4.12)

Therefore, the more unfair the easier to estimate $p_H$ accurately (because $V = p_H - p_H^2$), but, for example, 10% unfairness is not very easy to detect.

Perhaps it is fun to simulate the experiments described above computationally, but if you wish to do it, wait till you learn the large deviation.

4.7 Monte Carlo integration
Let us consider the problem of numerically evaluating a high-dimensional integral (the Monte-Carlo integration method):

$$ I = \int_{0}^{1} dx_1 \cdots \int_{0}^{1} dx_{1000} f(x_1, \ldots, x_{1000}). $$

(4.13)

If we wish to sample (only) two values for each variable, we need to evaluate the function at $2^{1000} \sim 10^{300}$ points (you should remember $2^{10} \simeq 10^3$). Such sampling is
of course impossible.

This integral can be interpreted as the average of $f$ over a 1000 dimensional unit hypercube:

$$I = \frac{\int_0^1 dx_1 \cdots \int_0^1 dx_{1000} f(x_1, \cdots, x_{1000})}{\int_0^1 dx_1 \cdots \int_0^1 dx_{1000}}.$$  \hspace{1cm} (4.14)

Therefore, randomly sampling the points $r_n$ in the hypercube, we can obtain

$$I = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(r_n).$$  \hspace{1cm} (4.15)

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 can readily know the answer from (4.11): $V(f(X_1))10^{\frac{7}{77}}$. The variance of the value of $f$ is of order $\max |f|^2$, a constant. Compare this number with $10^{300}$ above and appreciate the power of randomness. This is the principle of the Monte Carlo integration. Notice that the computational cost does not depend on the dimension of the integral.

How fast or slow the convergence of this method is may be felt from the estimation of $\pi$ by peppering a disk:

http://demonstrations.wolfram.com/MonteCarloEstimateForPi/

\textsuperscript{77}Here, the inequality gives a sufficiently safe estimate. In practice, a smaller number of samples might be OK.
5 Maxwell’s distribution

Summary
* Maxwell’s distribution of the particle velocity is derived.
* How to calculate Gaussian integral and averages is explained.
* Boltzmann factor $e^{-\beta U}$ is derived and used to obtain Maxwell’s distribution (again).
* Although slightly advanced, try to understand how to use the $\delta$-function (how to use practically: Appendices 5A).

Key words
Density distribution function, Maxwell’s distribution, Boltzmann factor, Gaussian integral, $\delta$-function

What you should be able to do
* Be able to use distribution functions (to estimate expectation values). Recall the use of indicators (see Appendix 5A).
* Understand how to estimate the mean-free path (or the idea of ‘sweep volume’).
* The molecular speed in a gas is of the same order of the sound speed in it.
* Be able to explain intuitively what the $\delta$-function is. Also you would better be able to use it systematically to compute various distribution functions (say, the distribution of the kinetic energy).

In this lecture, some computational techniques (not mere tricks but the ones practically useful) will be explained. Even if these explanations are not understood, the physics of the topics would be understandable, so the technical explanations are all in fine letters. If you do not read them, you would not encounter big difficulties (but you’d better browse through them at least). Before going to our new topics, let us review the law of large numbers (LLN).

5.1 LLN review: no fluctuation of internal energy

As we will learn the totality of mechanical energy of a macroscopic system is called

---

[78]Very strictly speaking, we do not know yet how big $N$ is, although we have already discussed this informally in 1.5. Empirically at this point, perhaps we could measure the temperature fluctuation of the system and guess how many ‘statistically independent parts’ make up the system.
the internal energy. The law of large numbers (Section 3) and the equipartition of energy (Section 2) imply that for an ideal gas the internal energy does not fluctuate macroscopically. Let us see why; this is a review of the last two Sections.

Let us take an ideal gas consisting of \( N \) particles in an isolated volume. We have demonstrated that all the particles have the same average kinetic energy \( \frac{3}{2} k_B T \) (equipartition of kinetic energy). Hence, the law of large numbers tells us

\[
P \left( \left| \frac{1}{N} \sum_{n=1}^{N} \frac{1}{2} m v_n^2 - \frac{3}{2} k_B T \right| > \varepsilon \right) < \frac{V(\varepsilon)}{\varepsilon^2 N},
\]

(5.1)

where \( V(\varepsilon) \) is the variance of the kinetic energy of each particle. Or, since \( E = \sum (1/2) m v_n^2 \) is the internal energy for the (monatomic) ideal gas,

\[
P \left( \left| \frac{E}{N} - \frac{3}{2} k_B T \right| > \varepsilon \right) < \frac{V(\varepsilon)}{\varepsilon^2 N}.
\]

(5.2)

This implies that the probability that the internal energy lies in the following range

\[
N(3k_B T/2 - \varepsilon) < E < N(3k_B T/2 + \varepsilon)
\]

(5.3)

is larger than \( 1 - V(\varepsilon)/\varepsilon^2 N \). Here, in practice, \( N \varepsilon \) need not be microscopic. We are dealing with a macroscopic body with \( N \sim 10^{20} \). Practically, we may need \( E \) within a 1\% error (\( \varepsilon = 0.01(3k_B T/2) \)). Notice that \( V(\varepsilon) \) is of the same order of \( (k_B T)^2 \). This implies that the probability to observe \( E \) in the range of (5.3) is larger than \( 1 - c/(0.01)^2 N \), where \( c = V(\varepsilon)/(3k_B T/2)^2 \) and is a constant of order unity. Thus, we have realized that ‘surely’ \( E \) is constant within 1\%. Actually, even if we increase the observation accuracy to \( 10^{-5}\% \) still the situation does not change very much for a 1 liter gas.

Let us proceed to a new topic.

\textbf{5.2 Density distribution function}

To make the kinetic theory quantitative, we must know the probability of a particle to assume various velocities. For the velocity of a particle to be \( v \) exactly is obviously with probability zero.
In the present case, the sample space is \( \Omega = \{ \mathbf{v} \mid v_x, v_y, v_z \in \mathbb{R} \} = \mathbb{R}^3 \). Thus, we need a probability measure\(^{80}\) \( P \) defined on \( \Omega \). In the present case, for a set \( A \subset \Omega \), \( P(A) \to 0 \) as volume\(^{81}\) of \( A \to 0 \),\(^{82}\) so we may define the probability density; symbolically (see Fig. 5.1),\(^{83}\)

\[
f(\mathbf{v}) = \frac{P(d\tau(\mathbf{v}))}{d\tau(\mathbf{v})},
\]

(5.4)

where \( d\tau(\mathbf{v}) \) is the volume element (of the 3-space) around \( \mathbf{v} \), which may be written as \( d^3\mathbf{v} = dv_x dv_y dv_z \). Here, its volume is also denoted by the same symbol \( d\tau(\mathbf{v}) \).

The probability \( P(A) \) of event \( A \subset \Omega \) may be expressed as

\[
P(A) = \int_A d^3\mathbf{v} f(\mathbf{v}).
\]

(5.5)

\(^{79}\)3-real vector space; generally, ‘\( n \)-object’ implies \( n \)-dimensional object.

\(^{80}\)’Probability measure’: You may interpret this as the correct use of the math terminology, if you know what measure is. Here, however, you may simply understand it informally as the usual probability, noting that the concept of probability is just the ‘volume of our confidence’ as we discussed in Lect 3, esp., around 3.1 and 3.3.

\(^{81}\)This is the actual volume of \( A \) as a subset of 3-space, which mathematicians call the Lebesgue measure of \( A \).

\(^{82}\)Mathematicians say that the probability measure \( P \) is absolutely continuous.

\(^{83}\)Actually, this notation is mathematically justified as the Radon-Nikodym derivative.
5.3 Maxwell’s derivation of Maxwell’s distribution function

In his “Illustrations of the dynamical theory of gases” (1860) Maxwell introduced the density distribution function $f(v)$ of the velocity of gas particles.

Let us follow Maxwell’s logic. He assumed that in equilibrium\footnote{What is equilibrium? It is a state reached by a gas isolated in a box sufficiently long after its preparation. There is no macroscopic flow in it and the gas is spatially uniform and time-independent (if observed on the macroscale).} orthogonal components of the velocity of particles are statistically independent. This implies (why? recall 3.9) that we may write

$$f(v) = \phi_x(v_x)\phi_y(v_y)\phi_z(v_z), \quad (5.6)$$

where $\phi_x$, etc., are density distribution functions for individual components. Maxwell also assumed the isotropy, so $f$ is a function of $v^2 \equiv |v|^2$, $f(v) \equiv F(v^2)$ and $\phi_x$, etc., do not depend on the suffixes specifying the coordinates: $\psi(s^2) \equiv \phi_x(s) = \cdots$. Therefore,

$$F(x + y + z) = \psi(x)\psi(y)\psi(z). \quad (5.7)$$

Maxwell originally assumed the differentiability of the functions, but here we only assume that the density distribution function is continuous. Since we are interested in the functional form of the density distribution function, and the normalization constant can be determined later, let us assume $\psi(0) = 1$.\footnote{If you do not like this, simply set $\psi(0) = a > 0$, a constant. Then, instead of (5.9) you get $G(x + y) = G(x) + G(y) - b$, where $b = 3 \log a$. Defining $g = G - b$, we get $g(x + y) = g(x) + g(y)$, from which we obtain $F(x) \propto e^{cx}$.} Then,

$$F(x + y) = \psi(x)\psi(y) = F(x)F(y). \quad (5.8)$$

Let $G(x) = \log F(x)$. Then, (5.8) reads

$$G(x + y) = G(x) + G(y), \quad (5.9)$$

which is called the Cauchy functional equation, whose general solution is $G(x) = cx$, where $c$ is a constant, if we assume $G$ is continuous (or monotone) (as we see just below 5.4).

This implies $F(x) \propto e^{cx}$; remember that normalization constant must be determined. That is, we may write with a new constant $c > 0$

$$f(v) \propto e^{-cxv^2}. \quad (5.10)$$

We should not forget, however, that Maxwell actually did not like the above derivation that assumed statistical independence of three orthogonal directions. He rederived it later with a different logic. Note that the above logic cannot work in 1-dimensional space.
5.4 Cauchy’s functional equation

\( G(2x) = 2G(x) \) is immediately obtained from (5.9). Repeating this, we get \( G(nx) = nG(x) \) for positive integer \( n \). This implies \( nG(1/n) = G(1) \) or \( G(1/n) = G(1)/n \). Therefore, \( G(m/n) = mG(1/n) = (m/n)G(1) \) for positive integers \( m \) and \( n \). Also, \( G(0) = 2G(0) \), so \( G(0) = 0 \). This implies \( G(x) = -G(x) \). Thus, we have demonstrated that for \( q \in \mathbb{Q} \) (rational numbers) \( G(q) = cq \), where \( c = G(1) \) is a constant. Since we assume \( G \) to be continuous (because \( F \) is positive and continuous), \( G(x) = cx \) must hold for any real \( x \).

5.5 Gaussian integral

We must compute the normalization constant and \( c \) in (5.10). An easy way (the easiest way?) to compute the normalization constant is the following elegant method.

Since the integral is positive, let us compute the square of what we want:

\[
\left( \int_{-\infty}^{\infty} dx \, e^{-x^2/2\sigma^2} \right)^2 = \int_{-\infty}^{\infty} dx \, e^{-x^2/2\sigma^2} \int_{-\infty}^{\infty} dy \, e^{-y^2/2\sigma^2} = \int_{\mathbb{R}^2} dx dy \, e^{-(x^2+y^2)/2\sigma^2}. \tag{5.11}
\]

Now, we go to the polar coordinates \((x, y) \rightarrow (r, \theta)\):

\[
\int_{\mathbb{R}^2} dx dy \, e^{-(x^2+y^2)/2\sigma^2} = 2\pi \int_{0}^{\infty} e^{-r^2/2\sigma^2} r dr = 2\pi \int_{0}^{\infty} dz \, e^{-z/\sigma^2} = 2\pi \sigma^2. \tag{5.12}
\]

Hence,

\[
\int_{-\infty}^{\infty} dx \, e^{-x^2/2\sigma^2} = \sqrt{2\pi} \sigma. \tag{5.13}
\]

5.6 Gaussian density distribution function

The Gaussian density distribution function \( g(x) \) generally has the following form:

\[
g(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-m)^2/2\sigma^2}, \tag{5.14}
\]

where \( E(x) = m \) and \( V(x) = \sigma^2 \) (recall the notations: 3.11, 3.12). Thus, we can know a Gaussian (density) distribution function, if we know its expectation value and variance.\(^{86}\)

When the density distribution function for a quantity \( x \) is given by the Gaussian form (5.14), we say the quantity obeys a Gaussian distribution or, simply, is Gaussian.

---

\(^{86}\)We will discuss the general multivariate Gaussian distribution in Lecture 18. We will learn that if we know the expectation value and the covariance matrix, we can fix the multivariate Gaussian distribution.
5.7 Maxwell’s distribution

Since \( \langle v_x \rangle = 0 \) and \( V(v_x) = (2/m)(k_B T/2) = k_B T/m \) (thanks to the equipartition of kinetic energy, 2.11), and since our distribution is Gaussian, we have

\[
\phi(v_x) = \sqrt{\frac{m}{2\pi k_B T}} e^{-mv_x^2/2k_B T}.
\]

That is,

\[
f(v) = \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2/2k_B T}.
\]

This is Maxwell’s distribution function. You must be able to compute various probabilities and expectation values with the aid of Maxwell’s distribution function (see Appendix 5A also).

Up to this point we discussed in 3-space, but the general \( d \)-Maxwell distribution should be obtained easily.

Exercise. The mode speed \( v_M \) of the particles is the value of \( v = \|v\| \) for which the probability for the particles to have the speed between \( v \) and \( v + dv \) becomes the largest. What can you expect for \( \langle v \rangle /v_M \) in the \( d \to \infty \) limit, where \( \langle v \rangle \) is the average speed (the average of \( \|v\| \))? [See Appendix 5A]

5.8 Generating function

At this juncture, let us practice a basic trick. You must be able to compute the expectation value of \( e^{\alpha x} \) for a Gaussian distribution, where \( \alpha \) is generally a complex number:

\[
\langle e^{\alpha x} \rangle = \frac{1}{\sqrt{2\pi \sigma}} \int_{-\infty}^{\infty} dx e^{\alpha x-(x-m)^2/2\sigma^2}.
\]

If \( \alpha = -s \), it is the Laplace transform of the distribution function; if \( \alpha = ik \), where \( k \) is real, \( \langle e^{ikx} \rangle \) is the Fourier transform of the density distribution, and is called the generating function.\(^87\)

The standard trick to compute this integral is to complete the square as follows (to make the form \( A(x-B)^2 + C)\):

\[
\alpha x - \frac{1}{2\sigma^2}(x-m)^2 = \alpha(x-m)+\alpha m - \frac{1}{2\sigma^2}(x-m)^2 = -\frac{1}{2\sigma^2}(x-m-\sigma^2 \alpha)^2 + \frac{\sigma^2 \alpha^2}{2} + \alpha m.
\]

\(^87\)http://www.yoono.org/ApplicableMath/ApplicableMath.html, Chapters 32 and 33 give a practical summary.
Therefore, (5.17) can be rewritten as
\[
\langle e^{\alpha x} \rangle = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} dx \, e^{-(x-m-\sigma^2\alpha^2)/2}\sigma^2 + \sigma^2\alpha^2/2 + am. \tag{5.19}
\]
We shift the integration range by \( m + \sigma^2\alpha \) (or you introduce a new integration variable \( x' = x - m - \sigma^2\alpha \) and rewrite the integral). Then, the integration just gives the normalization factor, so
\[
\langle e^{\alpha x} \rangle = e^{\sigma^2\alpha^2/2 + am}. \tag{5.20}
\]
Notice that
\[
\frac{d}{d\alpha} \langle e^{\alpha x} \rangle \bigg|_{\alpha=0} = E(x) = m, \tag{5.21}
\]
and
\[
\frac{d^2}{d\alpha^2} \langle e^{\alpha(x-m)} \rangle \bigg|_{\alpha=0} = V(x) = \sigma^2. \tag{5.22}
\]

5.9 Daniel Bernoulli revisited
Using Maxwell’s distribution function, let us review Bernoulli’s kinetic interpretation of pressure 2.10.

The (kinetic interpretation of) pressure on the wall is the average momentum (impulse) given to the wall per unit time and area. Consider the wall perpendicular to the \( x \)-axis (just as was in the elementary discussion in Lecture 2). Then, the number of particles with its \( x \)-component of the velocity being between \( v_x \) and \( v_x + dv_x \) that can impinge on the unit area on the wall per unit time is given by
\[
_nv_x dv_x \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_z f(v), \tag{5.23}
\]
where \( n \) is the number density (= number of particles in unit volume) of the gas molecules. Each particle gives the momentum \( 2mv_x \) upon hitting the wall, so
\[
P = \int_{v_x \geq 0} dv \, 2mv_x^2 f(v) = \int dv \, mnv_x^2 f(v) = \frac{1}{3}mn\langle v^2 \rangle, \tag{5.24}
\]
where we have used the symmetry \( f(v) = f(-v) \), and the isotropy, \( \langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle = (1/3)\langle v_x^2 + v_y^2 + v_z^2 \rangle \). Or,
\[
PV = \frac{2}{3}N\langle K \rangle, \tag{5.25}
\]
where \( K \) is the kinetic energy of the single gas particle, and \( N \) is the number of particles in the volume \( V \). Thus, Bernoulli’s equation has been derived once more (but rather mechanically).
5.10 Important Remark: $N$ was not known
Although the Boyle-Charles law is obtained, this does not tell us anything about molecules, because we do not know $N$. We cannot tell the mass of the particle, either. Remember that $k_B$ was not known when the kinetic theory of gases was being developed.

A notable point is that with empirical results that can be obtained from strictly equilibrium studies, we cannot tell anything about molecules, even about their existence.

Remember that the law of combining volumes for chemical reactions (2.8 (iii)) crucial to demonstrate the particular nature of chemical substances is about (often violent) irreversible processes from the reactants to the products. If we make a tiny hole on the wall of the container, we could make a molecular beam, so in principle, we can measure $\langle K \rangle$. However, this is a study of a system far away from equilibrium.

The (root-mean-square) speed of the molecules can be computed correctly, however, because we only need $PV/Nm$. Thus, in 1857 Clausius calculated the speed of molecules at $0^\circ C$: oxygen 461m/s, nitrogen 492m/s, and hydrogen 1,844m/s.\footnote{This is in his paper, “The Nature of the Motion which we call Heat” (Annalen der Physik, 100, 353 (1857); English translation in Phil Mag 14, 108 (1857)). Millet’s The Gleaners is this year.} Notice, that these speeds are not very different from the sound speeds in respective gases.\footnote{Sound speeds (in the standard state): oxygen 317 m/s, nitrogen 337 m/s, hydrogen 1270 m/s. As you see, this is about 2/3 of the molecular speed. For ideal gases, this ratio is exact.}

Let us derive Maxwell’s distribution in a more elementary fashion, following Feynman. We start with the derivation of the Boltzmann factor. This first part should be clearly understood. The rest may require some maturity of the reader; elementary approaches need not be very simple, so do not worry to much even if you do not understand this derivation of Maxwell’s distribution when you read it for the first time.

5.11 The Boltzmann factor
Take a vertical column of gas with cross-section $A$ in the gravitational field just around us.

Consider the force balance on the slice between heights $h$ and $h + dh$ of a cylinder. If $n$ is the number density and $m$ the mass of the molecule, we have as a force balance
equation with the aid of $P = nk_B T$

$$n \times Adh \times mg = -A(P + dP) + AP = -AdP = -Ak_B Tdn$$  

(5.26)

or

$$\frac{dn}{dh} = -\beta nmg,$$  

(5.27)

where $\beta = 1/k_B T$ (a standard abbreviation we will use often). That is,

$$n = n_0 e^{-\beta mg h}.$$  

(5.28)

This can describe the \textit{sedimentation equilibrium} of colloidal particles.

This equation suggests how the (relative) number of molecules depends on the potential energy difference. More generally, the same logic derives for conserved forces with potential $U$

$$n(r) = n(0)e^{-\beta[U(r) - U(0)]}.$$  

(5.29)

That is, the factor $e^{-\beta U}$ called the \textit{Boltzmann factor} tells us the ratio of particle densities (or the probabilities to find particles) at different locations, when there is a position dependent potential energy $U.$\footnote{This is true even if $U$ is extremely complicated. Thus, even if $U$ is not due to an external effect but due to other molecules in the system, this relation holds. This is also explained in the Feynman lectures. If you do not take the interactions with other molecules into account, then you cannot use the ideal gas law, but the actual equation of state for the fluid.}

\textit{Remark:} We have assumed that in equilibrium the temperature is everywhere the same independent of the height along the column. Is this correct? A particle going upward must lose its kinetic energy, so aren’t particles with lower temperatures at higher locations? According to this idea, since the number density on the lower side of a horizontal cross section is larger than that on the opposite side, such a number density distribution cannot be stably maintained. Thus, at least we realize that the temperature of the particles cannot be lower at higher locations. A more detailed balance argument can establish that the temperature must be uniform throughout the column in equilibrium.
5.12 Elementary derivation of Maxwell’s distribution

Consider a column of an ideal gas which is in equilibrium with gravity. Let \( n_{>u}(z) \) be the number of particles with \( v_z > u > 0 \) passing through height \( z \) upward per second (Fig. 5.3).

![Figure 5.3: If \( mu^2/2 = mgh \), then \( n_{>u}(0) = n_{>0}(h) \).](image)

Since stationarity of the distribution implies \( n_{>u}(0) = n_{>0}(h) \), if \( mgh = (1/2)mu^2 \),

\[
\frac{n_{>0}(h)}{n_{>0}(0)} = \frac{n_{>u}(0)}{n_{>0}(0)}.
\]  (5.30)

Since \( n_{>0}(h)/n_{>0}(0) = n(h)/n(0) \), we can use the Boltzmann factor just obtained:

\[
\frac{n_{>u}(0)}{n_{>0}(0)} = e^{-\beta mgh} = e^{-\beta mu^2/2}.
\]  (5.31)

The derivation is for the case without collisions, but since we have only to track energies that are conserved, collisions do not change the situation at all.

Let \( n(0, u) \) be the number density of particles at height 0 with the \( z \)-component of the velocity being \( u \). Notice that more numerous faster particles pass height 0 than slower ones, so we must take care of the speed in the \( z \)-direction:

\[
n_{>u}(0) = \int_u^\infty u n(0, u) \, du \propto e^{-\beta mu^2/2}.
\]  (5.32)

Thus, differentiating this equation, we obtain \( n(0, u) \propto e^{-\beta mu^2/2} \).

Now, we wish to go to a technical topic that will be crucial in more advanced physics, and also make many calculations about distributions quite mechanical (no special wisdom or insight is needed!).

**How to read the rest of this lecture:** You must clearly understand 5.13-5.15: what the \( \delta \)-function is, and how it is related to the density distribution function. The rest may be skipped for the first time reading.

5.13 What average gives the density distribution?

We have learned in 3.13 that the probability \( P(A) \) is given by the expectation value
of the indicator $\chi_A$ of event $A$:

$$P(A) = \langle \chi_A \rangle_P.$$  \hfill (5.33)

Therefore, we may write (5.4) formally as (suffix $P$ is omitted)

$$f(u) = \langle \frac{\chi_{d\tau(u)}}{d\tau(u)} \rangle.$$  \hfill (5.34)

Here, $\chi_{d\tau(u)}$ is the indicator of the volume element $d\tau(u)$ around $u$.

5.14 Let us introduce $\delta$-function

How does the quantity $\chi_{d\tau(u)}/d\tau(u)$ we formally obtained look like as a function of $v$? See Fig. 5.4 for the two-dimensional case. In the figure the actually infinitesimal volume element $d\tau(u)$ is assumed to be a tiny finite square for the illustration sake.

![Figure 5.4](attachment:image.png)

Figure 5.4: $\chi_{d\tau(u)}/d\tau(v)$, where the indicator of the volume element is concentrated around $u$. Its value is $1/d\tau(u)$ on the volume element around $u$, but otherwise zero.

Informally,

$$\frac{\chi_{d\tau(u)}(v)}{d\tau} = \begin{cases} 1/d\tau = \infty, & \text{if } v \in d\tau(u), \\ 0, & \text{otherwise.} \end{cases}$$  \hfill (5.35)

Here, $\infty$ appears because $d\tau$ is infinitesimally small. Then, following Dirac, let us introduce the $\delta$-function (delta function) (in the $d$-dimensional space) concentrated at $u$ as (here $v$ is assumed to be the running variable, and $u$ is a fixed constant vector. See Fig. 5.4.)

$$\delta(v - u) = 0 \text{ if } v \neq u,$$  \hfill (5.36)

$$\delta(v - u) d^d v = 1 \text{ if the volume element } d^d v \text{ is at } u.$$  \hfill (5.37)
This implies that for any continuous function $\varphi$ of $v$

$$
\int \varphi(v)\delta(v-u)\,dv = \varphi(u).
$$

(5.38)

Intuitively, as a function of $v$ you can imagine $\delta(v-u)$ as an infinitely thin but infinitely long needle at $v = u$ whose total volume is unity.

### 5.15 Formal expression of density distribution

Suppose we know the probability measure $P$ for a vector $v$. Then the density distribution function $f(v)$ for $v = u$ may be written as

$$
f(u) = \langle \delta(v-u) \rangle_P = \langle \delta(u-v) \rangle_P,
$$

(5.39)

where suffix $P$ denotes the distribution for $v$. Notice that ‘$\delta$-function’ may be regarded as an even function (see Fig. 5.6). You must keep in mind with respect to what variable you are taking the expectation value. In the above case it is with respect to $v$ and $u$ is a mere parameter while you compute the expectation with respect to $v$.

### 5.16 Distribution function: 1D case

Let us look at the 1D case in more detail.

Let $X(\omega)$ be a real stochastic variable defined on a probability space $(\Omega, P)$. The probability, denoted here by $F(x)$, for the event $X \leq x$ is called the **distribution function** of $X$:

$$
F(x) = P(X \leq x).
$$

(5.40)

The indicator of the event $X \leq x$ may be written in terms of the unit step function $\Theta$:

$$
\Theta(y) = \begin{cases} 
1, & \text{for } y \geq 0, \\
0, & \text{for } y < 0
\end{cases}
$$

(5.41)

as $\Theta(x-X)$, so we have (Fig. 5.5)

$$
F(x) = \langle \Theta(x-X) \rangle.
$$

(5.42)

Here, the average is over $X$.

---

\(^{91}\)In the following expression, if you wish to be very explicit you can write $P(X \leq x) = P(\{\omega | X(\omega) \leq x, \omega \in \Omega\})$.

\(^{92}\)If you wish to be more explicit, you should write $\langle \Theta(x-X) \rangle = \langle \Theta(x-X(\omega)) \rangle$. The average is over random events $\omega \in \Omega$. 

60
5.17 Density distribution function: 1D case
The derivative $f(x)$ of $F(x)$ is called the density distribution function of $X$. Using
the fundamental theorem of calculus, we can write as
$$F(x) = \int_{-\infty}^{x} f(x) dx.$$  \hspace{1cm} (5.43)
Physicists often call $f$ the distribution function.$^{93}$

If we can ‘differentiate’ the step function $\Theta(x)$, we should be able to write the
density distribution function in a neat form. An analogy to (5.39) suggests
$$f(x) = \langle \delta(x - X) \rangle_P,$$ \hspace{1cm} (5.44)
using the delta function adapted to 1D. Therefore, $\delta(x) = \Theta'(x)$ must be true.

5.18 1D $\delta$-function
This is indeed true. The step function $\Theta(x)$ is constant except at $x = 0$, so its
‘derivative’ $\delta(x) = 0$ if $x \neq 0$. However, it cannot be 0 everywhere. It is impossible
to ‘differentiate a vertical wall,’ but we intuitively see it to be $+\infty$; $\delta(0) = \infty$! The
area between $\delta(x)$ and the $x$-axis must be unity, because $\Theta$ jumps exactly by 1 at
$x = 0$. You may intuitively imagine an infinitely long needle at $x = 0$ whose ‘area’
is $0 \times \infty = \text{unity}$ as noted in the $d$-dimensional general case above.$^{94}$ Our intuition
just explained may be illustrated as Fig. 5.6.

$^{93}$Therefore, to distinguish it from the more standard usage of the terminology $F$ is sometimes
called the cumulative distribution function.
$^{94}$This is total nonsense! This cannot be mathematics! However, we can almost completely
rationalize our ‘intuitive picture’ with the aid of the theory of distributions due to Schwartz. Fig.
5.6 Right tells us that $0 \times \infty = 1$ is not so absurd.
Figure 5.6: The derivatives (Right) of increasingly steep cliffs (Left). Colors are correspondent. Imagining such a limiting process, we can understand the \( \delta \)-function as the derivative of the step function. Notice that the areas below the graphs in Right are always unity.

Therefore, for any continuous function \( \varphi(x) \) on the real line (see (5.38))

\[
\int_{-\infty}^{+\infty} \varphi(x) \delta(x - a) dx = \varphi(a). \tag{5.45}
\]

Thus, the Fourier transformation of \( \delta(x) \) is

\[
\int_{-\infty}^{+\infty} e^{ikx} \delta(x) dx = 1. \tag{5.46}
\]

Inverse Fourier transforming this, we get

\[
\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} dk. \tag{5.47}
\]

This is perhaps the most useful identity in theoretical physics.

The delta function is quite important in (theoretical) physics. A practical summary (as well as a short introduction to the general theory of such ‘hyperfunctions’ may be found in


In the following, a ‘practical minimum’ is outlined.

5.19 Key formula 1 to utilize the \( \delta \)-function

Let \( \alpha (\neq 0) \) be a real number. We have the following important identity:

\[
\delta(\alpha(x - a)) = \frac{1}{|\alpha|} \delta(x - a). \tag{5.48}
\]
This may be shown as
\[ \int_{-\infty}^{\infty} dx \delta(\alpha x - \alpha a) \varphi(x) = \int_{-\infty}^{\infty} d(y/|\alpha|) \delta(y - \alpha a) \varphi(y/\alpha) = \frac{1}{|\alpha|} \varphi(a), \] (5.49)

### 5.20 Key formula 2 to utilize the \( \delta \)-function

How about \( \delta(g(x)) \) for a differentiable function \( g \)? Suppose \( x_0 \) is a simple real zero of \( g(x) \). That is, assume \( g(x) \simeq g'(x_0)(x - x_0) \) near \( x_0 \). Then, (5.48) suggests that near \( x = x_0 \), \( \delta(g(x)) \) must be
\[ \delta(g(x)) = \frac{1}{|g'(x_0)|} \delta(x - x_0). \] (5.50)

There might be more then one simple real zeros of \( g(x) \) (i.e., \( g(r_1) = g(r_2) = \cdots = 0 \)). Thus, we obtain
\[ \delta(g(x)) = \sum_i \frac{1}{|g'(r_i)|} \delta(x - r_i) = \frac{1}{|g'(x)|} \sum_i \delta(x - r_i), \] (5.51)
where the summation is over all the simple real zeros of \( g \). For example, assuming \( a > 0 \)
\[ \delta(x^2 - a^2) = \frac{1}{2a} \left[ \delta(x - a) + \delta(x + a) \right]. \] (5.52)

### 5.21 Some practice

Let’s have some practice:
\[ \int_0^{10} dx \delta(x - \pi/6) \sin x = \frac{1}{2}, \] (5.53)
\[ \int_{-1}^{1} dx \delta(3x) \cos x = \frac{1}{3}, \] (5.54)
\[ \int_{-1}^{1} dx \delta(\pi - 6x) \cos x = \frac{1}{6} \cos \frac{\pi}{6} = \frac{\sqrt{3}}{12}, \] (5.55)
\[ \int_{-1}^{4} dx \delta(\pi - 6x) \cos x = 0, \] (5.56)
\[ \int_{-\infty}^{\infty} dx \delta(x^2 - 3x - 10)x^3 = \frac{1}{7} (125 - 8) = \frac{117}{7}. \] (5.57)
5.22 Kinetic energy density distribution of 2D ideal gas

Since we have seen the energy distribution of a 2D gas in a demo simulation, let us obtain its formula, and demonstrate that it is independent of the particle mass. Our starting point is a formal expression of the density distribution:

$$f(K) = \langle \delta(mv^2/2 - K) \rangle,$$

(5.58)

where the average is over the 2D Maxwell distribution. Therefore, we must cook

$$f(K) = \int dv_x dv_y \left( \frac{m}{2\pi k_B T} \right) e^{-mv^2/2k_B T} \delta((m/2)(v_x^2 + v_y^2) - K).$$

(5.59)

Rewrite this into a 1D integral with the aid of the polar coordinates $(v, \theta)$. We know the distribution must be isotropic, we can integrate over the direction $\theta$ to get $2\pi$, so

$$f(K) = 2\pi \left( \frac{m}{2\pi k_B T} \right) \int dv v \delta(mv^2/2 - K) e^{-mv^2/2k_B T}$$

(5.60)

$$= \pi \left( \frac{m}{2\pi k_B T} \right) \int dv^2 \delta(mv^2/2 - K) e^{-mv^2/2k_B T}$$

(5.61)

$$= \left( \frac{m}{2k_B T} \right) \int dx \delta(mx/2 - K) e^{-mx/2k_B T}.$$  

(5.62)

There are various ways to go further, but the easiest is to introduce $z = mx/2$ and to rewrite the above formula as

$$f(K) = \left( \frac{1}{k_B T} \right) \int dz \delta(z - K) e^{-z/k_B T} = \left( \frac{1}{k_B T} \right) e^{-K/k_B T}.$$

(5.63)

This is just what we guessed.
Appendix 5A: More calculations related to the Maxwell distribution

These days you may say no analytical muscle is needed thanks to Matlab, Mathematica, Maple, etc. However, wise use of these softwares requires good pattern recognition capability and strategic thinking. Such skills may be largely innate (genetic) but still can be nurtured considerably by practice.

The density distribution function of the particle velocity $v$ of mass $m$ at temperature $T$ is given by the following Maxwell distribution function:

$$f(v) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-mv^2/2k_BT}.$$  (5.64)

In this Appendix various expectation values and related distribution functions are studied.

(1) The (density) distribution function $F(u)$ of $u = |v|$. As noted in (??), we have

$$F(u) = \langle \delta(u - |v|) \rangle$$

or

$$F(u) = \langle \delta(u - |v|) \rangle = \int_{|v|} \delta(u - |v|) \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-mv^2/2k_BT} \, dv.$$  (5.65)

From this we can compute the mode (= the most probable value) of $v = 'u$ that maximizes $F(u)$' (use the logarithmic derivative for simplicity):

$$\frac{d}{du} \left(2 \log u - mu^2/2k_BT\right) = 0,$$  (5.66)

so $\sqrt{2k_BT/m}$ is the mode speed (typical speed).

The average speed $\langle v \rangle$ is

$$\langle v \rangle = \int_0^\infty dv \, F(v) v = \int_0^\infty \sqrt{\frac{2}{\pi}} \left(\frac{m}{k_BT}\right)^{3/2} v^2 e^{-mv^2/2k_BT} \, dv.$$  (5.67)
\[
= 2 \int_0^\infty \sqrt{\frac{2}{\pi}} \left( \frac{m}{k_B T} \right)^{3/2} x e^{-mx/k_B T} dx \quad \text{(here } x = v^2/2) \quad (5.72)
\]
To compute this integral, we use the following relations:
\[
\int_0^\infty x e^{-ax} dx = -\frac{d}{da} \int_0^\infty e^{-ax} dx = -\frac{d}{da} \frac{1}{a} = \frac{1}{a^2}. \quad (5.73)
\]
Thus,
\[
\langle v \rangle = 2 \sqrt{\frac{2}{\pi}} \left( \frac{m}{k_B T} \right)^{3/2} \left( \frac{k_B T}{m} \right)^2 = \sqrt{\frac{8k_B T}{\pi m}}. \quad (5.74)
\]

(2) Energy distribution: for \( E = mv^2/2 \) find its density distribution function \( F(E) \).
Just as above,
\[
F(E) = \langle \delta(E - mv^2/2) \rangle = \int_0^\infty 4\pi v^2 dv \delta(E - mv^2/2) \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2/2k_B T}. \quad (5.75)
\]
Then, we use (5.51) to cook the delta function with a nontrivial variable:
\[
\delta(g(x) - E) = \delta(x - x_E) \frac{1}{|g'(x_E)|}, \quad (5.76)
\]
where \( g(x_E) = E \). Applying this to our case, we have
\[
\delta(E - mv^2/2) = \delta(v - \sqrt{2E/m}) \frac{1}{\sqrt{2mE}}. \quad (5.77)
\]
Therefore, (you can use the result of (1) as well)
\[
F(E) = \int_0^\infty 4\pi v^2 dv \delta(v - \sqrt{2E/m}) \frac{1}{\sqrt{2mE}} \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2/2k_B T} \quad (5.78)
\]
\[
= 4\pi \frac{2E}{m} \frac{1}{\sqrt{2mE}} \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-E/k_B T} \quad (5.79)
\]
\[
= 2\sqrt{\frac{E}{\pi}} \left( \frac{1}{k_B T} \right)^{3/2} e^{-E/k_B T}. \quad (5.80)
\]

(3) Relative velocity distribution
Let us obtain the root-mean square relative velocity \( \mathbf{w} \) of two particles 1 and 2 with mass \( m \) in an ideal gas at \( T \).
Using the delta function technique, we can write the density distribution function \( F(w) \) for the relative velocity \( w = v_1 - v_2 \) as

\[
F(w) = \langle \delta(w - (v_1 - v_2)) \rangle_{v_1,v_2}
= \left( \frac{m}{2\pi k_B T} \right)^3 \int d^3v_1 \int d^3v_2 \delta(w - (v_1 - v_2))e^{-mv_1^2/2k_BT-mv_2^2/2k_BT}
= \left( \frac{m}{2\pi k_B T} \right)^3 \int d^3v_1 e^{-mv_1^2/2k_BT-m(w-v_1)^2/2k_BT}.
\]

(5.81)

(5.82)

(5.83)

The integration over \( v_1 \) is performed with the aid of completion of square (as explained in the lecture):

\[
mv_1^2 + m(w - v_1)^2 = 2mv_1^2 + mw^2 - 2mw \cdot v_1 = 2m(v_1 - w/2)^2 + mw^2 - (m/2)w^2 = 2m(v_1 - w/2)^2 + (m/2)w^2.
\]

(5.84)

(5.85)

(5.86)

Thus, we obtain

\[
F(w) = \left( \frac{m}{4\pi k_B T} \right)^{3/2} e^{-mw^2/4k_BT}.
\]

(5.87)

This means

\[
\langle w^2 \rangle = \frac{6k_BT}{m}.
\]

(5.88)

Check that the answer agrees with the result obtained by the equipartition of energy (and statistical independence of two particles).

(4) Root mean square velocity: we know an easy way, but here let us compute \( \langle v^2 \rangle \) honestly as

\[
\langle v^2 \rangle = \int_{v \in \mathbb{R}^3} d^3v \, v^2 \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2/2k_BT}.
\]

(5.89)

It is convenient to use the polar coordinate system with \( v = |v| \). Then,

\[
\langle v^2 \rangle = \int_{0}^{\infty} 4\pi v^2 dv \, v^2 \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2/2k_BT}.
\]

(5.90)

Thus, the following integral is needed:

\[
I(s, \alpha) = \int_{0}^{\infty} dv \, v^s e^{-\alpha v^2},
\]

(5.91)
which can be written with the aid of the Γ-function, but we do not go into it. For example,

\[ I(4, \alpha) = \frac{d^2}{d\alpha^2} \int_0^\infty dv \, e^{-\alpha v^2} = \frac{d^2}{d\alpha^2} I(0, \alpha) = \frac{d^2}{d\alpha^2} \frac{1}{2} \sqrt{\frac{\pi}{\alpha}} = \frac{3}{8} \frac{\sqrt{\pi}}{\alpha^{5/2}}. \]  

(5.92)

We know

\[ \langle v^2 \rangle = \frac{I(4, \alpha)}{I(2, \alpha)} = \frac{3}{2} \frac{1}{\alpha} = \frac{3 k_B T}{m}. \]  

(5.93)
Q5-1. Consider $D$-dimensional ideal gas ($D > 1$). That is, the velocity is a $D$-vector $\mathbf{v} = (v_1, v_2, \cdots, v_D)$, and the kinetic energy of a single particle is given by $(m/2)v^2 = (m/2)(v_1^2 + v_2^2 + \cdots + v_D^2)$, where $m$ is the mass of the particle. [Here, our approach is elementary, but you can use the $\delta$-function technique already explained in Appendix 4B. If the problem becomes more complicated as the next problem, the $\delta$-function is much easier than the elementary approach.]

1) Write down the $D$-dimensional Maxwell distribution function (i.e., find the density distribution function of the velocity in $D$-space).

2) What is the most probable speed $v_D$ in $D$-space? That is, what is the mode speed (the speed for which the density distribution function for the speed is maximal)?

3) What is the ratio of $v_D$ obtained in (2) and the root-mean-square velocity in the $D \to \infty$ limit?

4) Is your result in (3) consistent with the law of large numbers?

Soln.

1) As you can guess from the 3D result, we have only to multiply $D$ 1D results as (If you do not like mere guessing, you can go back to Maxwell’s proof. You will realize that the proof boils down to $F(x + y) = F(x)F(y)$ just as in 3D.)

$$f(\mathbf{v}) = \left( \frac{m}{2\pi k_B T} \right)^{D/2} e^{-m|\mathbf{v}|^2/2k_B T}.$$ 

2) We need the density distribution function $F(v)$ of the speed $v = |\mathbf{v}|$.

$$F(u) = \langle \delta(u - |\mathbf{v}|) \rangle = \int d^D \mathbf{v} \delta(u - |\mathbf{v}|) \left( \frac{m}{2\pi k_B T} \right)^{D/2} e^{-m|\mathbf{v}|^2/2k_B T}$$

Notice that to use the rule of the computation of the integral containing a $\delta$-function, the independent variable in the $\delta$-function (in our case $|\mathbf{v}|$) must be the integration variable. Thus we must convert the $\delta$-function or convert the integration variable (in our case, we convert the integration variable from $\mathbf{v}$ to $|\mathbf{v}|$).

Now we should go to the polar coordinate system in $D$-space:

$$d^D \mathbf{v} = S_{D-1} u^{D-1} dv,$$

where $S_{D-1}$ is the volume of the $D-1$-unit sphere (corresponding to $4\pi$ in 3D). This actual form may be found in my grad course lecture notes (there is a clever way to compute it as explained there), but we do not need its explicit form. Integrating over the velocity, we get

$$F(u) = S_{D-1} \left( \frac{m}{2\pi k_B T} \right)^{D/2} u^{D-1} e^{-mu^2/2k_B T}.$$
This is the density distribution function for the speed. To find its peak, we have only to maximize \( \frac{mu}{k_B T} - (D - 1) \frac{1}{u} = 0 \):

\[
\frac{mu}{k_B T} - (D - 1) \frac{1}{u} = 0,
\]

Or \( u^2 = (D - 1)k_B T/m \). That is,

\[ v_D = \sqrt{\frac{(D - 1)k_B T}{m}}. \]

(3) The equipartition of kinetic energy tells us that

\[
\langle \frac{1}{2} m v^2 \rangle = \frac{mD}{2} \langle v_x^2 \rangle = \frac{D}{2} k_B T.
\]

That is, the root-mean-square velocity is \( \sqrt{Dk_B T/m} \). The ratio obviously converges to unity.

(4) The LLN tells us that for any positive \( \varepsilon \)

\[
P \left( \left| \frac{v^2}{D} - \frac{k_B T}{m} \right| > \varepsilon \right) < \frac{V(v^2)}{\varepsilon^2 D}.
\]

That is, the probability for

\[
D \left( \frac{k_B T}{m} - \varepsilon \right) < v^2 < D \left( \frac{k_B T}{m} + \varepsilon \right)
\]

is asymptotically unity as \( D \rightarrow \infty \) for any \( \varepsilon > 0 \). Since \( F(u) \) has peak(s) in this range, the most probable value should be within this range as well. That is, the mode speed is forced to agree with the root-mean square average.

**Q5-2** [Density distribution of relative velocity]

(1) There are two particles 1 and 2 in an equilibrium pure ideal gas. Write down the simultaneous density distribution function \( f(v_1, v_2) \) of their velocities \( v_1 \) and \( v_2 \). You may assume the temperature of the gas is \( T \), and the mass of the individual particles is \( m \).

(2) Now, introduce the velocity \( V \) of the center of mass of these two particles and the relative velocity \( w = v_1 - v_2 \). Write down the simultaneous density distribution function \( g(w, V) \) of \( w \) and \( V \).

(3) Find the density distribution function of \( w \). Compute \( \langle |w| \rangle \) and compare it with
the root-mean square of $w$ (i.e., $\sqrt{\langle w^2 \rangle}$).

**Soln.**

1. Maxwell’s (density) distribution function $f(v)$ implies

$$P(dv) = f(v)dv,$$  
(5.94)

where $dv$ is the volume element of the velocity space. Since we know two particles are statistically independent,

$$P(dv_1, dv_2) = P(dv_1)P(dv_2),$$  
(5.95)

the density $f(v_1, v_2)$ must be a product of two Maxwellian distributions. Therefore,

$$f(v_1, v_2) = \left(\frac{m}{2\pi k_B T}\right)^3 e^{-m(v_1^2 + v_2^2)/2k_B T}.$$  
(5.96)

2. $v_1 = V + w/2$ and $v_2 = V - w/2$, so

$$v_1^2 + v_2^2 = 2V^2 + \frac{1}{2}w^2.$$  
(5.97)

Since we are computing the density distribution, we must demand

$$f(v_1, v_2) dv_1 dv_2 = g(V, w) dV dw,$$  
(5.98)

or

$$f(v_1, v_2) \frac{\partial (v_1, v_2)}{\partial (V, w)} = g(V, w).$$  
(5.99)

The Jacobian appearing in the above formula is unity, so

$$g(V, w) = \left(\frac{m}{2\pi k_B T}\right)^3 e^{-(2mV^2 + (m/2)w^2)/2k_B T}.$$  
(5.100)

Notice that the center of mass kinetic energy is $(1/2)(2m)V^2$, and the kinetic energy of the relative motion is $(1/2)(m/2)w^2$, where $m/2$ is the reduced mass. You can read them off from the above formula.

3. The marginal distribution $g(w)$ is obtained by integrating $V$ out, or simply splitting $g(V, w)$ using statistical independence of $V$ and $w$:

$$g(w) = \left(\frac{m}{4\pi k_B T}\right)^{3/2} e^{-mw^2/4k_B T}.$$  
(5.101)
\[ \langle |w| \rangle = \left( \frac{m}{4\pi k_B T} \right)^{3/2} \int dw |w| e^{-mw^2/4k_B T} = \left( \frac{m}{4\pi k_B T} \right)^{3/2} 4\pi \int_0^\infty w^3 e^{-mw^2/4k_B T} dw \]

(5.102)

The integral can be calculated analytically:

\[ 4\pi \int_0^\infty w^3 e^{-mw^2/4k_B T} dw = 2\pi \left( \frac{4k_B T}{m} \right)^2 . \]

(5.103)

Therefore,

\[ \langle |w| \rangle = 2\pi \left( \frac{4k_B T}{m} \right)^2 \left( \frac{m}{4\pi k_B T} \right)^{3/2} \]

(5.104)

On the other hand

\[ \langle w^2 \rangle = \left( \frac{m}{4\pi k_B T} \right)^{3/2} 4\pi \int_0^\infty w^4 e^{-mw^2/4k_B T} dw = \left( \frac{m}{4\pi k_B T} \right)^{3/2} 48\pi^{3/2} \left( \frac{k_B T}{m} \right)^{5/2} = 6k_B T \]

(5.105)

or \( \sqrt{\langle w^2 \rangle} = 3\sqrt{2}\sqrt{k_B T/m} \), which we will use in the next Lecture to estimate the mean free path. This must be larger than \( \langle w \rangle \), because the variance of \( w = |w| \) (i.e, \( \langle w^2 \rangle - \langle w \rangle^2 \)) must be positive.

**Q5-3.** Let us obtain the root-mean square relative velocity of two molecules with different masses \( m \) and \( M \) in an equilibrium gas at temperature \( T \).

(1) Using the delta function trick, we can write the density distribution function \( F(w) \) for the relative velocity \( w \) as

\[ F(w) = \left( \frac{m}{2\pi k_B T} \right)^{3/2} \left( \frac{M}{2\pi k_B T} \right)^{3/2} \int d^3v_1 \int d^3v_2 \delta(w-(v_1-v_2))e^{-mv_1^2/2k_B T-Mv_2^2/2k_B T}. \]

(5.106)

Perform the integration over \( v_2 \).

(2) Then, perform the integration over \( v_1 \) to obtain \( F(w) \).

(3) Find \( \langle w^2 \rangle \) and check that the answer agrees with the result obtained by the equipartition of energy.

**Soln.**

(1) This is straightforward. By inspection, we get

\[ F(w) = \left( \frac{m}{2\pi k_B T} \right)^{3/2} \left( \frac{M}{2\pi k_B T} \right)^{3/2} \int d^3v_1 e^{-mv_1^2/2k_B T-M(w-v_1)^2/2k_B T}. \]

(2) To perform the Gaussian integral we use the trick to complete the square:

\[ mv_1^2 + M(w-v_1)^2 = (m + M)v_1^2 + Mw^2 - 2Mwv_1 \]

(5.107)
Thus, we obtain

\[ F(w) = \left( \frac{mM}{2(m + M)\pi k_B T} \right)^{3/2} e^{-\frac{mM}{m + M} w^2 / 2k_B T}. \]  

(5.110)

Notice that the appearance of the reduced mass is quite natural.

(3) The expectation value can be read off from the formula as

\[ \langle w^2 \rangle = 3 \frac{m + M}{mm} k_B T. \]

The result is of course consistent with the elementary results as follows:

\[ \langle w^2 \rangle = \langle v_1^2 \rangle + \langle v_2^2 \rangle = 3 \frac{k_B T}{m} + 3 \frac{k_B T}{M} = 3 \frac{m + M}{mm} k_B T. \]

Q5-4. We know, in equilibrium, the mean kinetic energy \( E \) is given by the equipartition of energy. What is the probability of a particle to have the kinetic energy more than the average kinetic energy \( 3k_B T / 2 \) (in equilibrium)?

Soln.

We know the density distribution of \( E \) in Appendix 4B, so we can use it, but here we proceed step by step.

\[ P \equiv P(E \geq 3k_B T / 2) = \langle \chi_{\{v:v^2 \geq 3k_B T / m\}} \rangle = \int_{\{v:v^2 \geq 3k_B T / m\}} f(v) dv, \]  

(5.111)

where \( f \) is Maxwell’s distribution. Therefore,

\[ P = 4\pi \int_{\sqrt{3k_B T / m}}^{\infty} \left( \frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2 / 2k_B T} v^2 dv \]  

(5.112)

\[ = \frac{4}{\sqrt{\pi}} \int_{\sqrt{3/2}}^{\infty} e^{-x^2} x^2 dx \simeq 0.39. \]  

(5.113)

This should be less than \( 1/2 \) because we can expect very high energy but rare particles.

Q5-5. There are \( N \) particles which can be in one of the two states A and B. State
A has a potential energy about $U = 0.2k_B T$ higher than B. How many particles $N$ do you need to estimate $U$ within 1% (relative) error from a single measurement of the occupation probability of state A (i.e., $N_A/N$) in equilibrium at temperature $T$ (according to the usual law of large numbers estimate)? [In this question, you must assume your ‘failure tolerance level’, e.g., a larger than $\varepsilon$ error once in 100 observations, or once in 500 observations, etc. Choose your tolerance level.]

**Soln.**

This is a LLN problem. $p_A = N_A/N$ is the empirical probability you can measure. We know $N_A/N_B = e^{-U/k_BT} \simeq e^{-0.2} = 0.8187$, and $N = N_A + N_B$

$$p_A = 1/(1 + e^{\beta U}) \simeq 0.45 \quad (5.114)$$

LLN tells us for any $\varepsilon (> 0)$, if $P_A$ is the true probability of state A (from which we may compute accurate $U$)

$$P(|p_A - P_A| > \varepsilon) < \frac{V(p_A)}{\varepsilon^2 N}, \quad (5.115)$$

where $V(p_A)$ is the variance of $p_A$: $V(p_A) = P_A(1 - P_A) \simeq P_A(1 - p_A) = 0.248$. We tolerate the relative error of 1% in $U$ (i.e., $\Delta U/U \sim 0.01$). The tolerated error is estimated as

$$|\Delta p_A| = p_A^2 e^{\beta U} \beta U (\Delta U/U) = 0.45^2 \times e^{0.2} \times 0.2 \times 0.01 = 0.000494. \quad (5.116)$$

Therefore, $\varepsilon = 0.0005$. Thus, the error bound in the above formula reads

$$\frac{V(p_A)}{\varepsilon^2 N} = \frac{0.248}{(0.0005)^2 N}. \quad (5.117)$$

If we allow one failure in 1000 observations, this must be less than 1/1000. Therefore,

$$N > 248/(0.0005)^2 = 1. \times 10^9. \quad (5.118)$$
6 Mean free path and transport phenomena

Summary
* Clausius introduced the concept of mean free path.
* Linear transport phenomena are outlined. Fluxes are proportional to (−)gradients of the density fields.
* Transport coefficients in the gas phase can be estimated with the aid of elementary kinetic theory.
* Maxwell used shear viscosity and the van der Waals equation of state to estimate Avogadro's constant and the molecular size for the first time.

Key words
Mean free path, linear transport phenomena, density, flux, gradient, divergence, conservation law, Laplacian, transport coefficient, diffusion, diffusion coefficient, shear viscosity, heat conductivity

What you should be able to do
* This lecture is a bit complicated for those who have never encountered partial differential equations such as the diffusion equation. Those who have no problem with the very rudimentary portion of partial differential equations should pay attention to the following entries. Those who feel this Lecture a bit too much should understand the concepts such as the mean-free path, density, flux, gradient and divergence at least intuitively, and try to understand the flow of Maxwell’s log.
* Rudimentary vector analysis should be reviewed (gradient, divergence, Laplacian; you must explain their intuitive meanings).
* Understand how to handle the averages of vector components.
* You should be able to understand how to derive the partial differential equation describing the conservation law.
* Recognize that there are some relations among transport coefficients; dimensional analysis is useful.

6.1 Mean free path
Dutch meteorologist C. H. D. Buys-Ballot (1817-1890)\textsuperscript{95} noticed that if the molecules

\textsuperscript{95}who noticed the Buys-Ballot law: In the Northern Hemisphere, if a person stands with his
of gases really moved that fast as Clausius estimated (see 5.10), the mixing of gases by diffusion should have been much faster than we observed it to be.

The (first half of the) following YouTube video about diffusion demonstrates the point:

http://www.youtube.com/watch?v=H7QsDs8ZRM1.\textsuperscript{96}

Upon this criticism, Clausius (1858\textsuperscript{97}) realized that the gas molecules have large enough diameters so a molecule cannot move very far without colliding with another one. In this way Clausius defined a new parameter called the \textit{mean free path} $\ell$ of gas that describes the average distance a molecule can run between two consecutive collisions. We can obtain it with the idea of ‘swept volume’ by a particle (see Fig. 6.1). The moving molecule sweeps a cylinder (‘swept volume’) of radius $d$ (\textit{=} the diameter of the molecule). If this volume does not contain any center of mass of other molecules, no intermolecular collision occurs. If it contains one, there is a collision. Therefore, if the swept volume $\times n \sim 1$, where $n$ is the number density, the height of the cylinder must be the ‘mean free path’ length. Hence, we guess

$$\ell = 1/n\pi d^2, \quad (6.1)$$

if all other particles are fixed in space, where $\pi d^2$ is the cross-section of the swept volume.

Actually, all the molecules are moving. When they collide, the average relative speed must be the relevant velocity, which is $\sqrt{2}$ times the mean velocity. That is,

\textsuperscript{96}However, you must take into account that the demo is affected by gravity, because Br\textsubscript{2} is far heavier than air.

\textsuperscript{97}1858: the Lincoln-Douglas debate, the Government of India Act. Planck (~1947) was born. However, the most important event was that the idea of natural selection was officially published by Darwin and Wallace. Physicists should recognize that Boltzmann called the 19th century the century of Darwin (not of Maxwell) (see E. Broda, \textit{Ludwig Boltzmann, Mensch-Physiker-Philosoph} (F Deuticke, 1955) Part III).}
the molecule collides $\sqrt{2}$ times more often than the case where all other molecules are fixed in space. Therefore,

$$\ell = \frac{1}{\sqrt{2\pi nd^2}}$$

must be the true mean free path length.\footnote{If you sit on one particle and observe other particles, their mean speed is, on the average, $\sqrt{2}$ times the actual mean speed of the particles (relative to the coordinates fixed to the ground). Thus, collisions become $\sqrt{2}$ times more frequent than when other particles are still.}

### 6.2 Why transport phenomena matter

Clausius did not have any method to estimate $\ell$. However, as we can expect from the criticism by Buys-Ballot, if we could study the so-called transport phenomena, there is a hope to determine $\ell$. This is a step toward estimating $N$. This was exactly what Maxwell accomplished through his study of viscosity (= transport of momentum) of gases. He performed the first realistic estimation of Avogadro’s number. To understand what he accomplished, we must know a bit about transport phenomena.

### 6.3 What is a (linear) transport phenomenon?

Suppose a macroscopic system is not far away from equilibrium. The system may be spatially nonuniform, but is macroscopically only gently so. For example, the number density of the molecules in the system may not be spatially constant and may be described as a number-density field $n(t, r)$, where $t$ is time and $r$ is the spatial position vector.

If there is a gentle spatial nonuniformity in some physical quantity $X$,\footnote{In transport phenomena, we are interested in ‘extensive quantities.’ We will learn what they are later.} there is a field of its density $\hat{x}(t, r)$. We can expect a flow of this physical quantity to reduce the nonuniformity. Thus, $X$ must be transported from one point to another. This is generally called the transport phenomenon. If $\partial \hat{x}/\partial t$ is a linear functional of $\hat{x}$,\footnote{$F$ being a ‘linear functional’ implies the following: $F(aX_1 + bX_2) = aF(X_1) + bF(X_2)$} we say the transport phenomenon is linear.

### 6.4 Density

Let $X$ be a physical quantity carried by molecules. Its density around space-time
point \((t, r)\) may be expressed as

\[
\hat{x}(t, r) = \frac{\sum_{i \in d\tau(r)} x_i}{d\tau(r)},
\]

(6.3)

where \(x_i\) is the amount of \(X\) carried by the \(i\)th molecule whose spatial location is \(r_i\) at time \(t\). Here, \(d\tau(r)\) indicates the volume element around \(r\), which is very small from the macroscopic point of view, but it is huge from the microscopic molecular point of view. Its volume is also denoted by the same symbol \(d\tau(r)\). The summation on the numerator means that we calculate the summation over particles whose centers of mass are in \(d\tau(r)\). The law of large numbers tells us that \(\hat{x}(t, r)\) thus defined is not appreciably fluctuating, so we identify it with the density of \(X\) at (around) \(r\) at time \(t\).

6.5 Flux

To describe the flow of \(X\), we need the concept of flux. A flux \(J_X\) of \(X\) is a vector pointing in the direction of the flow, whose magnitude is the amount of the quantity going through the unit cross section per unit time (see Fig. 6.2). We may often write \(J_X\) to be the product of the density of \(X\) and the velocity of the underlying flow carrying it: \(J_X = \hat{x} v\).

![Figure 6.2: The flux vector \(J_X\) for the quantity \(X\) (here, its density is denoted by \(\hat{x}\)): its direction is the transport direction, and its magnitude is the flow rate: the quantity of \(X\) through the area \(A\) perpendicular to \(J_X\) (converted to the amount per unit area) per unit time.]

This expression may be microscopically written as

\[
J_X = \frac{\sum_{i \in d\tau(r)} x_i v_i}{d\tau(r)},
\]

(6.4)

If the transport phenomenon is linear, the flux \(J_X\) of \(X\) is proportional to the gradient of its density \(\text{grad} \, \hat{x}(r)\) (its direction is opposite; cf. Fig. 6.3; here time \(t\) is suppressed):

\[
J_X = -L \text{grad} \, \hat{x}(r),
\]

(6.5)

\(^{101}\)But still macroscopic in the sense that the number of particles in it is, say, \(10^{10}\). For an ordinary gas in the so-called standard state (at 1 atm-273K), its volume is about \(200 \times 10^{-6}\) mm\(^3\) \(\approx\) a cube with 60 \(\mu\)m edge. It may be a bit larger than our cells.

78
where $L$ is a positive constant called the transport coefficient.

\[ \hat{x}(r) \]

\[ \text{grad } \hat{x} \]

\[ J_x \]

Figure 6.3: Gentle nonuniformity causes linear transport phenomena. The gradient vector $\text{grad } \hat{x}$ points in the direction of increasing density $\hat{x}$ (darker region), so the flux driven by the gradient points in the $-\text{grad } \hat{x}$ direction.

The gradient of $\hat{x}$ is the following vector:

\[ \text{grad } \hat{x} \equiv \nabla \hat{x} = \frac{\partial \hat{x}}{\partial x} e_x + \frac{\partial \hat{x}}{\partial y} e_y + \frac{\partial \hat{x}}{\partial z} e_z, \tag{6.6} \]

where $e_k$ is the directional vector (unit vector) in the $k$-axis direction. That is, componentwisely,

\[ \text{grad } \hat{x} = \left( \frac{\partial \hat{x}}{\partial x}, \frac{\partial \hat{x}}{\partial y}, \frac{\partial \hat{x}}{\partial z} \right). \tag{6.7} \]

$\nabla$ is an operator called nabla (usually it is read as ‘del’) and may be understood as the following vector:

\[ \nabla = e_x \frac{\partial}{\partial x} + e_y \frac{\partial}{\partial y} + e_z \frac{\partial}{\partial z} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \tag{6.8} \]

$\nabla \hat{x}$ may be understood as the product of a vector $\nabla$ and a scalar $\hat{x}$ (needless to say, you cannot change the order of this product, since $\nabla$ is operating on $\hat{x}$).

Intuitively, you can imagine a landscape with altitude $X$ given as a function of the position, and then a vector pointing the steepest ascending direction at a location $r$ with its size give by the slope of the landscape along the vector at $r$. The vector is ‘$\text{grad } X$’ at $r$.

### 6.6 Divergence

If $X$ is conserved, the amount of change of this quantity at a given position must be equal to the net influx of $X$ to that position. Therefore, if we introduce an operator div (read as ‘divergence’) that allows us to compute the net output of $X$ from a point based on the flux $J_X$ at the position, the conservation law for $X$ may be expressed as

\[ \frac{\partial \hat{x}(r)}{\partial t} = -\text{div } J_X(r). \tag{6.9} \]
Here, ‘div’ is out-going quantity, so − is put. The divergence $\text{div } J_X$ of the flux $J_X$ of $X$ at point $P$ (the total amount of output per unit volume per unit time) may be defined as:

$$\text{div } J_X = \lim_{V \to P} \frac{\int_{\partial V} J_X \cdot dS}{\int_V d\tau}.$$  \hspace{1cm} (6.10)

Here, $\lim_{V \to P}$ implies the limit along the sequence of nested (singly connected) volumes $V$ converging to point $P$ (Fig. 6.4) with its surface denoted by $\partial V$. $dS$ is the surface area element, whose direction is the outward normal direction, and whose magnitude (area) is $dS$. Thus, the numerator on the right-hand side is the total amount of $X$ going out of the volume $V$ in unit time.

If we use the Cartesian coordinate system,

$$\lim_{V \to P} \frac{\int_{\partial V} J_X \cdot dS}{\int_V d\tau} = \frac{[J_x(x + dx, y, z) - J_x(x, y, z)]dydz + [J_y(x, y + dy, z) - J_y(x, y, z)]dxdy + [J_z(x, y, z + dz) - J_z(x, y, z)]dxdy}{dxdydz}.$$ \hspace{1cm} (6.11)

That is,

$$\text{div } J_X = \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = \nabla \cdot J_X.$$ \hspace{1cm} (6.12)

The rightmost expression implies that divergence can be formally written as the scalar product of $\nabla$ and the flux vector. If you need a review of vector analysis, go to, e.g., Section 2.C. of


6.7 Local expression of conservation law

Suppose the density $\hat{x}$ is conserved. The total amount of $X$ coming into the volume

---

102 $\partial A$ is the standard notation for the boundary of the set $A$. 

element $d\tau = dx dy dz$, that is, $-\text{div } J_X dx dy dz$ must be the increase of $X$ in it. Therefore, we have
\[
\frac{\partial \hat{x}}{\partial t} dx dy dz = -\text{div } J_X dx dy dz,
\tag{6.13}
\]
that is, the conservation equation (6.9) has been derived. If $X$ can be produced with the rate $\sigma$ per unit volume (say, due to a chemical reaction), (6.9) is modified to the following general conservation law with production:
\[
\frac{\partial \hat{x}}{\partial t} = -\text{div } J_X + \sigma.
\tag{6.14}
\]

6.8 Diffusion equation
Let us first study the simplest linear transport phenomenon: the diffusion of particles. We know the number of particles is conserved without any chemical reaction. Therefore, if $J$ is the number flux, (6.9) is
\[
\frac{\partial n}{\partial t} = -\text{div } J.
\tag{6.15}
\]
We assume linear transport of particles (called Fick’s law)
\[
J = -D \text{ grad } n,
\tag{6.16}
\]
where $D$ is the diffusion coefficient. Combining these two, we get
\[
\frac{\partial n(t, r)}{\partial t} = -\text{div } (-D \text{ grad } n(t, r)) = D \nabla \cdot (\nabla n(t, r)).
\tag{6.17}
\]
This is the conservation law for the particle number density called the diffusion equation. Introducing the Laplacian $\Delta$ as
\[
\Delta = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},
\tag{6.18}
\]
the diffusion equation reads
\[
\frac{\partial n(t, r)}{\partial t} = D \Delta n(t, r).
\tag{6.19}
\]
6.9 The meaning of Laplacian

If you understand the meaning of the Laplacian, you will feel the diffusion equation very natural. Let us consider the 1d Laplacian. It is nothing but $\frac{d^2}{dx^2}$. If we compute the second derivative numerically, we use, for example, the following discretization

$$\frac{d^2 f(x)}{dx^2} \leftarrow \frac{f'(x + \Delta x/2) - f'(x - \Delta x/2)}{\Delta x} = \frac{1}{\Delta x} \left( \frac{f(x + \Delta x) - f(x)}{\Delta x} - \frac{f(x) - f(x - \Delta x)}{\Delta x} \right),$$

(6.20)

so

$$\frac{d^2 f(x)}{dx^2} \propto \frac{f(x + \Delta x) + f(x - \Delta x)}{2} - f(x).$$

(6.21)

That is, $d^2 f/dx^2 \propto \text{local average of } f \text{ around } x' - f(x)$. You can confirm this conclusion, studying higher dimensional cases: generally, the Laplacian is an operator to compare the central value and the average value surrounding it. Thus, in the particle number diffusion the Laplacian computes the difference between the average $n$ surrounding $r$ and $n(r,t)$. If this is positive, the diffusion equation increases $n(r,t)$ in order for this quantity to catch up with the neighbors.

6.10 Intuitive computation of transport coefficient

To compute the transport coefficient for a quantity $X$ we need a microscopic description of $J_X$. Maxwell carried out this step fairly intuitively. Although it is hard to refine his argument quantitatively, as we will see soon, Maxwell’s rather crude argument allows a fairly realistic estimation of Avogadro’s constant.

Here, we start with a crude microscopic interpretation of a flux as the product of the flow velocity and the density (cf. 6.5). Basically, we understand that, on the average, a molecule brings the physical quantity of our interest adopted at the location of its latest collision to the location $r$ where it is now (see Fig. 6.5). If we write the ‘free vector’ (a displacement vector of a molecule between successive collisions) as $l_i$ for particle $i$, the last collision should have occurred at around $r - l_i$. No new collision occurs until the molecule arrives at the volume element around $r$, so the contribution of this molecule to the flux must be $x_i(r - l_i)v_i$, where $x_i(r - l_i)$
is quantity $X$ the molecule $i$ acquired at its last collision. We use (6.4) to get

$$J_X = \sum_{r_i \in d\tau(r)} x_i (r_i - l_i) v_i / d\tau(r),$$  \hspace{1cm} (6.22)

but usually it is further approximated as

$$J_X = \langle \hat{x}(r - l)v \rangle = \langle (l \cdot \nabla \hat{x}(r))v \rangle + \cdots,$$  \hspace{1cm} (6.23)

where the average over $l$ and $v$ is taken for molecules around $r$. $\langle \hat{x}(r)v \rangle = \hat{x}(r)\langle v \rangle = 0$ and vanishes, because $\hat{x}(r)$ is constant in the volume element.

To compute the second term in (6.23) let us consider

$$\langle (l \cdot A)v \rangle = \langle v(l \cdot A) \rangle = \left\langle v \sum_{i \in \{x, y, z\}} l_i A_i \right\rangle$$  \hspace{1cm} (6.24)

for an arbitrary vector $A$. $v$ and $l$ are parallel and each component of $v$ is statistically independent, so\textsuperscript{103*}

$$\langle v_i l_j \rangle \simeq \frac{1}{3} \overline{v} \ell \delta_{ij}.$$  \hspace{1cm} (6.25)

Here, $\overline{v}$ is the average speed of the particles and $\ell$ is the mean-free path. We have arrived at

$$\langle v(l \cdot A) \rangle = \frac{1}{3} \overline{v} \ell A.$$  \hspace{1cm} (6.26)

Thus, we have arrived at

$$J_X(r) = -\frac{1}{3} \overline{v} \ell \nabla \hat{x}(r).$$  \hspace{1cm} (6.27)

This is the general formula within Maxwell’s approach for the flux.

\textbf{6.11 Diffusion constant}

For Fick’s law (6.16) $\hat{x} = n$, so the diffusion constant is obtained as

$$D = \frac{1}{3} \overline{v} \ell,$$  \hspace{1cm} (6.28)

which may also be written as

$$D = \frac{\ell^2}{3 \tau},$$  \hspace{1cm} (6.29)

where $\tau$ is the mean free time $\tau = \ell / \overline{v}$.

As can be seen from the derivation above, the numerical factor $1/3$ is not quite a definitive number.\textsuperscript{104} The main message is that $D/\overline{v} \ell$ is a numerical factor of order

\textsuperscript{103*}Each component of $v$ is statistically independent, so it assumes $\pm$ independently. Both $\langle v \rangle$ and $\langle l \rangle$ are zero, so if $i \neq j$, $\langle v_i l_j \rangle = 0$. From the isotropy of the space we get $\langle v_i l_i \rangle = \langle v_2 l_2 \rangle = \langle v_3 l_3 \rangle = \langle v \cdot l \rangle / 3$, but $v$ and $l$ are parallel vectors, so their scalar product becomes the product of their lengths, i.e., mean speed $\overline{v}$ and mean free path $l$. Thus, we have arrived at (6.25).

\textsuperscript{104}Besides, what average speed $\overline{v}$ to use is not very clear.
unity. Then, this should be derivable dimensionally-analytically. Try this derivation (then read 6.15).

6.12 Shear viscosity

Suppose we have a shear flow with the velocity $V$ in the $x$-direction and the velocity gradient in the $z$-direction as shown in Fig. 6.6. To understand the decay of this velocity gradient we study the transport of the $x$-component of the momentum. Due to exchange of particles between positions with different $z$-coordinates, larger $V_x$ (or larger momentum density) and smaller $V_x$ layers mix and the gradient in the $z$ direction decays. This is the effect of shear viscosity.

![Figure 6.6: Shear flow](image)

Figure 6.6: Shear flow: We consider a macroscopic shear flow, so the gradient of $V$ must be microscopically (esp., on the scale of the mean-free path $l$) very small, but in the figure, it is exaggerated.

The derivation of (6.27) immediately tells us that if the transported density is $\hat{x}$, the corresponding flux reads

$$J_X(r) = -\frac{1}{3} \pi l \ell \text{grad} \hat{x}(r). \quad (6.30)$$

To apply this general formula to the quantity we are interested in, we must identify what $\hat{x}$ is. In our present case, it must be the $x$-component of the momentum density

$$\hat{x} = \sum_{d\tau} m v_x / d\tau, \quad (6.31)$$

where the summation in the numerator means to take the summation of all the $x$-components of the momentum of the particles in the volume element $d\tau$. Therefore, (here we assume the number density $n$ is uniform) thanks to the law of large numbers we expect only the expectation value is relevant, so

$$\hat{x}(r) = n m V_x(r) \quad (6.32)$$

is the right density to study. Therefore, (6.30) (or its $z$-component) reads

$$J_V = -\frac{1}{3} \pi l n m \frac{\partial V_x}{\partial z}, \quad (6.33)$$

84
where $J_V$ is the $z$-component of the ‘$x$-component momentum flux’.\footnote{In a more advanced course, we use a tensor.} Shear viscosity $\eta$ is defined by

$$J_V = -\eta \partial V_x / \partial z,$$  \hfill (6.34)

Comparing this with (6.33), we get the shear viscosity $\eta$:

$$\eta = \frac{1}{3} mn \bar{v} l.$$  \hfill (6.35)

With the already obtained estimate of $l$ (6.2) and $\bar{v} = \sqrt{8k_BT/\pi m}$, we obtain$^{106}\eta$:

$$\eta = \frac{2}{3d^2} \sqrt{\frac{m k_B T}{\pi^3}}.$$  \hfill (6.36)

This is independent of the density $n$ as noted by Maxwell. We generally expect that the viscosity increases with density, but in gases, higher densities imply shorter free paths or a shorter mixing distance (actually the mean free path length is $\propto 1/n$) and the expected density effect is cancelled. Also notice that the viscosity increases with temperature. Although this is contrary to the behavior we usually encounter in liquids, it is easy to understand because higher temperatures imply better mixing.

### 6.13 Maxwell estimated the molecular size$^{107}$

To establish the reality of atoms, we wish to determine the number of particles $N$ and their sizes $d$. Even if you could determine the mean-free path length, we can determine only the combination $Nd^2$.

In 1873$^{108}$ van der Waals (1837-1923) proposed his equation of state of imperfect gases:$^{109}$

$$P(V-V_0) = Nk_B T - \frac{\alpha}{V} (1 - V_0/V).$$  \hfill (6.37)

---

$^{105}$In a more advanced course, we use a tensor.

$^{106}$If we assume that the particle mass, the cross section ($d^2$) and the particle thermal velocity are only relevant quantities, dimensional analysis gives essentially this result. Even if we try to take the density of the gas into account, it automatically drops out of the formula. This independence was a bit of surprise. It is a good occasion to learn rudiments of dimensional analysis.

$^{107}$Actually, his first calculation of 1873 followed the method proposed by Loschmidt in 1865, who identified $\pi d^3/6$ as the volume per molecule in the liquid phase. Therefore, $(\pi d^3/6)/(1/n) = V_L/V_G$, where $V_L$ is the molar volume of the liquid phase and $V_G$ that of the gas phase. Thus, we obtain $d = 6\sqrt{2(V_L/V_G)} \ell$. Loschmidt estimated this as $8(V_L/V_G) \ell$, where $\ell$ was obtained from diffusion experiments (see (6.28)). Since we get $d$, we can count the number of molecules in $V_L$. Maxwell estimated $N_A \sim 4.3 \times 10^{23}$.

$^{108}$Maxwell’s A Treatise on Electricity and Magnetism was published this year, so was Jules Verne’s Around the World in Eighty Days.

$^{109}$We will study this equation in detail later (Section 25).
His basic idea is as follows (see Fig. 6.7): Since molecules are not point masses but have volumes, they cannot run everywhere they wish (at least they must avoid each other). However, if we collect all the volumes of the molecules at a corner of the container (its volume is $V_0$), then, the centers of mass of the molecules could freely move around in the ‘free volume’ $V - V_0$. Therefore, if we ignore the attractive interactions, the ‘hard-core’ gas would look like an ideal gas with a reduced volume:\footnote{As we will show in Section 25, his idea is correct in 1-space.}

\begin{equation}
    P(V - V_0) = N k_B T. \tag{6.38}
\end{equation}

The remaining part of the van der Waals equation is to take care of the attractive intermolecular forces. Thus, from $V_0 \simeq b N \pi d^3 / 6$, where $b$ is a geometrical constant of order unity, we can estimate the size of the molecules. Now, we know $N d^2$ and $N d^3$, so we can estimate $N$ and $d$. The method gives an estimate of Avogadro’s constant $N_A \simeq (4 \sim 6) \times 10^{23}$,\footnote{\cite{Definition of Avogadro’s constant} Avogadro’s constant is defined as the number of atoms in a 0.012 kg of $^{12}$C. The latest 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). Cf. P. Beker, “History and progress in the accurate determination of the Avogadro constant,” Rep Prog Phys 64 1945 (2001).}

### 6.14 Heat conductivity

The heat conductivity $\lambda$ is defined as

\begin{equation}
    \mathbf{J}_H = -\lambda \text{grad} T, \tag{6.39}
\end{equation}

where $\mathbf{J}_H$ is the heat flux (the thermal energy flux). The transported density $\dot{x}$ must be the thermal energy contained in the unit volume. Let us assume that the gas is a monatomic gas:

\begin{equation}
    \dot{x} = \sum_{d\tau} \frac{mv^2}{2} d\tau, \tag{6.40}
\end{equation}
\[
\dot{x}(r) = \frac{3}{2}nk_BT(r),
\] (6.41)

where \( T(r) \) is the temperature field.

(6.30) reads

\[
J_H(r) = -\frac{1}{3}v_{\parallel} \text{grad} \left( \frac{3}{2}nk_BT(r) \right) = -\frac{1}{2}nkB\overline{v} \text{grad} T(r).
\] (6.42)

Comparing this with (6.39), we obtain

\[
\lambda = \frac{1}{2}nkB\ell v.
\] (6.43)

Notice that \( \eta = nmD, \eta/\lambda = 2m/3nkB \) and \( \lambda/D = 3nkB/2 \). The last relation tells us \( \lambda = c_VD \), where \( c_V \) is the specific heat per molecule of gas under constant volume.\(^{112}\) Again, we should note that these relations do not tell us anything about the microscopic properties of the gas particles.

### 6.15 Dimensional analysis of transport coefficients\(^{113}\)

The dimension of a quantity \( X \) is usually denoted by \([X]\). The basic dimensions are represented by the following symbols: length \( L \), mass \( M \) and time \( T \). For example, \([d] = L\). To obtain the dimension of a quantity, go back to its definition. For example, \([D]\) is obtained from \( J = -D \text{ grad } n \) as follows. The particle number flux is the number of particles going through a unit area in unit time, so \([J] = 1/L^2T\), because the number of particles is dimensionless (a pure number). \([n] = 1/L^3\). Gradient is essentially differentiation with length, so \([\text{grad}] = 1/L\) (differentiation is something like division). Therefore, \([J] = 1/L^2T = [D]/L^4\), so \([D] = L^2/T\).

For \([\eta]\) let us go back to its definition: \( J_p = -\eta \text{ grad } v \), where \( J_p \) is the momentum flux, and \( v \) is the velocity. Since the dimension of momentum is \( ML/T \), \([J_p] = (ML/T)/L^2T = M/LT^2\), \([v] = L/T\), so \([\eta] = [J_p]L/[v] = M/LT\).

For \([\lambda]\) again let us go back to its definition \( J_H = -\lambda \text{ grad } T \) (in this formula \( T \) is temperature, so \( k_BT = E \) is energy). Therefore, \([E]/L^2T = [\lambda][E/k_B]/L\), so we obtain \([\lambda/k_B] = 1/LT\).

Thus, we obtain \([D/\eta] = L^3/M\), so \( \eta/D \sim mn \) is concluded. We get \([k_BD/\lambda] = L^3 = 1/[n]\), which gives \( \lambda \sim nk_BD \sim c_VD \). We also get \([k_B\eta/\lambda] = M\), which implies

\(^{112}\)\( c_V \) is the energy required to raise the temperature of the molecule by 1 K under constant volume.

\(^{113}\)\( \langle \text{Introduction to dimensional analysis} \rangle \) See, e.g., Appendix 3.5A of Oono, Y. (2013). \textit{The Nonlinear World}, Tokyo: Springer.
\[ \eta/\lambda \sim m/k_B. \] These are the relations mentioned above.

6.16 Significance of flux dependent on gradient

Due to collisions the particles cannot go straight for a long distance (actually, it is a zig-zag random walk as we will see in the next lecture). If there were no collision, the particles can move along their straight ‘ballistic’ trajectories, so the amount of ‘\( X \)' transported must be proportional to the difference of \( \dot{x} \) (not to the slope of \( \dot{x} \) called gradient as we learned for linear transport phenomena) irrespective of the distance over which transportation occurs. Thus, the flux proportional to the gradient is actually a clear sign of molecular collisions occurring on the microscopic scale.
Q6-1.
(1) What is the dimension of the heat conductivity (or thermal conductivity) divided by the Boltzmann constant $\lambda/k_B$? $\lambda$ is defined by
\[ J = -\lambda \text{ grad } T, \]
where $J$ is the heat flux (the flux of kinetic energy; transported kinetic energy per unit time through unit area).
(2) It is a natural guess that heat transport should be related to the amount of thermal energy carried by molecules and the transport rate of the molecules. The former may be represented by the specific heat per volume $c$ (times temperature) and the latter by diffusion constant $D$. What can dimensional analysis tell you about the relation among these quantities? We already discussed this in the lecture, but you must rederive the relation purely dimensionally-analytically.

Soln.
(1) 
\[ [J] = M(L/T)^2/L^2T = M/T^3 \]
\[ [\text{grad } k_B T] = M(L/T)^2/L = ML/T^2. \]
Therefore
\[ [\lambda/k_B] = (M/T^3)/(ML/T^2) = 1/LT. \] (6.44)
Thus $[\lambda/k_B] = 1/LT$. This is consistent with the unit of heat conductivity is W/m-K.
(2) We know $[D] = L^2/T$. The heat capacity is energy/(volume times temperature), so $[c/k_B] = 1/L^3$. Therefore, $[cD/k_B] = 1/LT$. Thus, $\lambda \propto cD$ may be concluded.

Q6-2. There is a pure gas which roughly obeys a van der Waals equation of state with the excluded volume $V_0 = 5.1 \times 10^{-5}$ m$^3$/mole. Note that
\[ V_0 = N_A \frac{4\pi}{3} d^3 = \frac{2\pi}{3} N_A d^3, \] (6.45)
where $N_A$ is Avogadro’s constant and $d$ is the diameter of the gas particle (atom or molecule spherically approximated).
(1) This gas has a density of 5.894 kg/m$^3$ under 1 atm at $T = 273$ K. What is the root-mean-square velocity of the gas particles for this gas?
(2) The diffusion coefficient was observed to be $D = 4.8 \times 10^{-6}$ m$^2$/s. Using the simple gas kinetic estimate of $D$ in the lecture notes (i.e., $D = lv/3$), obtain the mean free path length $l$. Here, you may identify $v$ with the root-mean-square velocity just computed in (i).
(3) Try to estimate Avogadro’s constant from the data given above.$^{114}$

Soln.
$^{114}$The data here is for xenon.
(1) \( \langle v^2 \rangle = 3P/\rho \), so \( \sqrt{3 \times 1.013 \times 10^5/5.89} = 227 \text{ m/s}. \)

(2) \( \ell = 3D/v = 6.34 \times 10^{-8} \text{ m}. \)

(3) Let \( V = 22.4 \times 10^{-3} \text{ m}^3 \) be the volume of this gas at 1 atm:

\[
N_A d^2 = V / \sqrt{2\pi \ell}, \quad N_A d^3 = 3V_0 / 2\pi.
\]

Therefore,

\[
d = \frac{3V_0 \sqrt{2\pi \ell}}{2\pi V} = \frac{3\ell V_0}{\sqrt{2V}},
\]

which is \( 3.067 \times 10^{-10} \text{ m} = 3.1 \text{ Å} \). A reasonable value. (van der Waals radius = 2.2 Å for xenon) and

\[
N_A = V / \sqrt{2\pi \ell d^2} = \frac{V}{\sqrt{2\pi \ell} 9\ell^2 V_0^2} = \frac{\sqrt{2} V^3}{9\pi \ell^3 V_0^2} = 8.466 \times 10^{23}.
\]
7 Brownian motion

Summary
* Brown found the universal motion called the Brownian motion.
* Einstein recognized the Brownian motion is due to thermal motion, and estimated \( k_B \) (or equivalently \( N_A \)).
* Langevin explains the Brownian motion in terms of the equation of motion with a noise term (called the Langevin equation).
* Brownian trajectories may be related to random walks and polymer chain conformations. \( \langle r^2 \rangle = 2dDt \).

Key words
Brownian motion, Langevin equation, mesoscopic, Fick’s law, Einstein’s relation, diffusion equation, Laplacian, Einstein-Stokes relation, dimensional analysis

What you should be able to do
* Be able to explain the key idea of the mesoscopic approach using Einstein’s Brownian motion theory as an example.
* Be able to derive Einstein’s relation.
* Be able to estimate the span of a random walk or a random chain polymer.

7.1 How mesoscopic particles behave
At the microscopic level molecules are colliding with the fellow molecules and are recoiling forever. What if the particle we observe is much bigger than the molecules surrounding it? The particles we can observe optically are about thousand times linearly as large as the molecules (Fig. 7.1). This means that the mass ratio is \( \sim 10^9 \). Thus, numerous small impulses are imparted to the big particle from the surrounding molecules. The law of large numbers tells us that the motion of the big particle must be extremely slow compared with the gas particles, and its motion is due to the ‘\( o[N] \)’ part of the law of large numbers (4.4). That is, we observe a typical mesoscopic scale motion, which we now call the Brownian motion.

7.2 Mr Brown discovered a universal motion (now) called the Brownian motion
Figure 7.1: Brownian particle (1 μm radius) vs molecules (1 nm radius); The sun/the earth ratio is about 110. The ratio of the radius of the orbit of the earth ( = 1 AU) and the radius of the sun is 109 (150 Gm vs. 1.4 Gm). This means the ratio of our cell (eukaryotic cell) and the molecule size is about the ratio of 1 AU and the size of the earth. Right from a nice site.

The Brownian motion was discovered in the summer of 1827\textsuperscript{115,116} by Robert Brown.\textsuperscript{117} We are usually given an impression that he simply observed the “Brownian motion.” However, he did a very careful and thorough research to establish the universal nature of the motion.

Since the particles for which Brown first observed the motion came from living cells (see Fig. 7.2), initially he thought that it was a vital phenomenon. Removing the effects of advection, evaporation, etc., carefully, he tested many flowers. Then, he tested old pollens in the British Museum (he was the (founding) director of the Botanical Division), and still found active particles. He conjectured that this was an organic effect, testing even coal with no exception found. This suggested him that not only vital but organic nature of the specimens were irrelevant. He then

\textsuperscript{115}[1827: Beethoven died in March; Democratic party was founded.]

\textsuperscript{116}The work was published the next year. See P Pearle, B Collett, K Bart, D Bilderback, D Newman, and S Samuels, “What Brown saw and you can too,” Am. J. Phys. 78, 1278 (2010).

\textsuperscript{117}⟨Who was Mr Brown?⟩ [See Cook, Banks, Humboldt, ..., Bates http://www.yoono.org/PST_Cambridge/Section9.html for a (historical) background] Robert Brown (1773-1858) was perhaps the greatest botanist (and a great microscopist; Alexander von Humboldt (1769-1859) called him ‘the glory of Great Britain’) in the first half of the 19th century. He wrote (1810) a classic of systematic botany describing the Australian flora, following his expedition (1801-5). He was the first to recognize the two major classes of seed plants (1827) [P. B. Tomlinson, “Rescuing Robert Brown—The Origins of Angio-Ovuly in Seed Cones of Conifers,” Bot. Rev. 78, 310 (2012)]. He recognized the nucleus of the cell and so named it (1831; the terminology was later imported by N. Bohr to atomic physics).

Before departing for his Beagle expedition (Dec., 1831-Oct., 1836), Charles Darwin (1809-1882) asked for Brown’s advice in 1831, buying a portable dissecting microscope recommended by Brown; after returning to England, Brown encouraged Darwin to visit him every Sunday morning. Later, Brown was regularly invited to parties at Darwin’s home.

The participants of the now historical Linnean Society meeting, where the theory of natural selection was first read (July 1, 1858), were there mainly to listen to Lyell reading the eulogy for Brown who died on June 10 and to praise his career. Cf. J. Browne, Charles Darwin, voyaging (Knopf, 1995), Charles Darwin, the power of place (Knopf, 2002); an authoritative biography of Charles Darwin.
tested numerous inorganic specimens (including a piece of Sphinx; he also roasted
his specimens).

Figure 7.2: The pollen tube Brown observed first was from *Clarkia pulchella* (flower reddish
purple, Oenotheraceae, Northwest US; the genus name commemorates Clark of the Lewis and
Clark expedition (1804-6)). He observed 1/4000-1/5000 in (0.5-0.6 $\mu$m) particles in the pollen
tube. From the quoted booklet: “the first plant examined proved in some respects remarkably well
adapted to the object in view. This plant was *Clarkia pulchella*, of which the grains of pollen, taken
from antherae full grown, but before bursting, were filled with particles of granules of unusually
large size, perhaps slightly flattened, and having rounded and equal extremities. While examining
the form of these particles immersed in water, I observed many of them very evidently in motion ...
” [USDA photo]

Let us watch some examples:
Nanoparticles in water:
http://www.youtube.com/watch?v=cDcprgWiQEY&feature=topics
Simulations
http://www.youtube.com/watch?v=PtYP8uoN0lk&feature=topics (excellent; comparison of
small and large particles)
This may be the best (again):
http://labs.minutelabs.io/Brownian-Motion/ may be fun.

7.3 General properties of Brownian motion
Curiously enough, there was no work published about Brownian motion between
1831 and 1857, but the phenomenon was well known. From 1850s new experimental
studies began by Gouy (1854-1926) and others. The established facts included (you
would find them very easy to understand in terms of molecular bombardment on
mesoscopic particles):

(1) Its trajectory is quite erratic without any tangent lines anywhere.
(2) Two Brownian particles are statistically independent even when they come within
their diameters.

93
(3) Smaller particles move more vigorously.
(4) The higher the temperature, the more vigorous the Brownian motion.
(5) The smaller the viscosity of the fluid medium, the more vigorous the motion.
(6) The motion never dies out.

etc.\textsuperscript{118}

In the 1860s there were experimentalists who clearly recognized that the motion was due to the impact of water molecules. \textit{Even Poincaré (1854-1912) mentioned this motion in 1900, but somehow no founding fathers of kinetic theory and statistical mechanics paid any attention to Brownian motion.}\textsuperscript{119}

Due to the bombardment of water molecules, the Brownian particle executes a zigzag motion, and eventually, say, its $x$-coordinate\textsuperscript{120} displaced as seen in Fig. 7.3; its source video is worth watching.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure7.3.png}
\caption{Displacement of Brownian particles along one coordinate; the $x$-axis is the time and the $y$ the position of various sample particles along a line. The rightmost figure schematically describe the density of the particles at the end of the journeys. [From the video quoted above]}
\end{figure}

As you see in Fig. 7.4 the salient feature of the Brownian displacement $\Delta \mathbf{r}$ is

$$\langle \Delta r^2 \rangle \propto t,$$ \hspace{1cm} (7.1)

where $\langle \rangle$ is the ensemble average (you repeat the experiment again and again or do many (mutually not interfering) experiments simultaneously, and average the results) and $t$ is time. The proportionality constant is related to (proportional to) the

\textsuperscript{118}What can you conclude from these observations and dimensional analysis?

\textsuperscript{119}According to H. Ezawa, they never expected the particle fluctuations large enough to be observable.

\textsuperscript{120}This figure illustrates originally a 1D Brownian motion, but it also illustrates the behavior of a particular component of the position vector of a single Brownian particle, since all the orthogonal coordinates of a 3D Brownian particle position vector is statistically independent.
7.4 Langevin’s explanation of the Brownian motion

Closely following Paul Langevin’s argument, let us demonstrate indeed \( \langle \Delta r^2 \rangle \propto t \).

Let us try to describe the motion of a Brownian particle classical mechanically. Let \( \mathbf{r} \) be its position vector, and \( m \) its mass. Newton’s equation of motion requires the forces acting on the particle. Since the particle is being hit ‘randomly,’ we expect a random force \( \mathbf{w} \) (whose direction and magnitude change incessantly and erratically) acting upon the particle. If the Brownian particle moves at a constant velocity \( \mathbf{v} \), then it would be hit by more particles of the medium on its front than on its back (imagine running in the rain). Therefore, it is natural to expect a force opposing the motion (i.e., drag) whose magnitude is proportional to the speed. Therefore, the equation of motion reads

\[
 m \frac{d^2 \mathbf{r}}{dt^2} = -\zeta \frac{d\mathbf{r}}{dt} + \mathbf{w},
\]

where \( \zeta \) is a positive constant describing the relation between the particle velocity and the resistive or frictional force the particle feels from the medium. Let us try to

make an equation for $r^2 = r \cdot r$ by scalar-multiplying $r$ to this equation. Since

$$r \cdot \frac{d^2 r}{dt^2} = \frac{d}{dt} \left( r \frac{dr}{dt} \right) - \left( \frac{dr}{dt} \right)^2 = \frac{d}{dt} \left( \frac{1}{2} \frac{dr^2}{dt} \right) - \left( \frac{dr}{dt} \right)^2,$$

we have

$$\frac{m}{2} \frac{d^2 r^2}{dt^2} - m \left( \frac{dr}{dt} \right)^2 = -\frac{\zeta}{2} \frac{dr^2}{dt} + w \cdot r.$$  \hspace{1cm} (7.4)

Let us ‘ensemble-average’ this equation. That is, we prepare many such Brownian particles and average the equations for them. Let us denote this averaging procedure by $\langle \rangle$. Since averaging procedure is linear and time-independent, we can exchange the order of differentiation and averaging. Thus, we obtain

$$\frac{m}{2} \frac{d^2 \langle r^2 \rangle}{dt^2} - m \langle \left( \frac{dr}{dt} \right)^2 \rangle = -\frac{\zeta}{2} \frac{d \langle r^2 \rangle}{dt} + \langle w \cdot r \rangle.$$  \hspace{1cm} (7.5)

Langevin says, “The average value of the term $w \cdot r$ is evidently null by reason of the irregularity of the complementary forces $w$.” Also, thanks to the equipartition of kinetic energy in equilibrium, the second term on the LHS is known:

$$\frac{1}{2} m \langle \left( \frac{dr}{dt} \right)^2 \rangle = \frac{3}{2} k_B T,$$

where $k_B$ is the Boltzmann constant and $T$ is the absolute temperature of the system.\(^{122}\)

If we introduce

$$z = \frac{d \langle r^2 \rangle}{dt},$$

(7.5) reads

$$\frac{m}{2} \frac{dz}{dt} + \frac{\zeta}{2} z = 3k_B T.$$  \hspace{1cm} (7.8)

Notice that this ‘3’ is the spatial dimensionality $d$. This implies after a sufficiently long time,\(^{123}\) the time derivative $dz/dt$ should vanish and $z = 6k_B T/\zeta$ (note that 6 here is $2d$), or

$$\langle r^2 \rangle = \frac{6k_B T}{\zeta} t = \frac{2d k_B T}{\zeta} t.$$  \hspace{1cm} (7.9)

\(^{122}\)Since it is very hard to measure the velocity of the Brownian particles, the equipartition of kinetic energy was hardly directly proved, so some people even doubted this.

\(^{123}\)which is actually a mesoscopic scale relaxation time: $\tau \approx m/\zeta$. This is very short for a macroscopic observer like us.
That is, the absolute value of the displacement during time $t$ is proportional to $\sqrt{t}$. See Fig. 7.3.

### 7.5 Relation to random walk
As we have seen, due to random bombardment by fluid particles a Brownian particle executes an erratic motion. Let $\Delta \mathbf{r}_i$ be the total displacement between time $(i-1)\tau$ and $i\tau$, where

![Figure 7.5: Actual observation results of a latex particle trajectory for 3.3 sec. Left: every 1/8000 sec; Right: every 1/30 sec. (Courtesy of Prof. Nishizaka of Gakushuin U)](image)

$\tau$ is a mesoscopic time scale which is very small from our point of view (say, 1 ms). Then, we may model the movement of the particle by a random walk (Fig. 7.5). After $n$ steps (after $t=n\tau$), the total displacement of the Brownian particle is given by

$$\mathbf{r}(t) = \Delta \mathbf{r}_1 + \Delta \mathbf{r}_2 + \cdots + \Delta \mathbf{r}_n.$$  \hfill (7.10)

Let us compute the mean square displacement:

$$\langle r^2 \rangle = \sum_i \langle \Delta r_i^2 \rangle + 2 \sum_{i<j} \langle \Delta r_i \cdot \Delta r_j \rangle.$$  \hfill (7.11)

Since the movement of the Brownian motion is uniform (e.g., throughout the duration of the motion the displacements are statistically the same), we may expect $\langle \Delta r_1^2 \rangle = \langle \Delta r_2^2 \rangle = \cdots$. Since there is no systematic direction to move into, $\langle \Delta \mathbf{r}_i \rangle = 0$. Since $\Delta \mathbf{r}_i$ are totally random (statistically independent), we expect that the average $\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle = \langle \Delta \mathbf{r}_i \rangle \cdot \langle \Delta \mathbf{r}_j \rangle = 0$ for $i \neq j$. Therefore, (7.11) implies

$$\langle r^2 \rangle = n\langle \Delta r_i^2 \rangle \propto t,$$  \hfill (7.12)
which is consistent with (7.9).

7.6 Let us look at 3D random walk samples
Let us observe 3D random walks (NN = 10000 steps), using the following R program (you can download R.app from CRAN https://cran.r-project.org):

```r
install.packages("ggplot2")
Probably, you are asked to choose a CRAN mirror site. Choose, say, USA(KS).

install.packages("rgl")
library(ggplot2)
library(rgl)

The actual program begins here. (If you wish to reset the shape, simply rerun the whole program by copying the following)

```r
NN <- 10000
m <- matrix(numeric(3*NN), ncol = 3)
for (i in 2:NN)
{
  q <- rnorm(3)
  m[i, ] <- m[i-1, ] + qn
}
df <- setNames(data.frame(m, seq(1, NN)),c("x", "y", "z"))
plot3d(df, xlim = c(-sqrt(NN), sqrt(NN)), ylim = c(-sqrt(NN), sqrt(NN)), zlim = c(-sqrt (NN), sqrt (NN)), type = "l")
```

7.7 Polymer chain as a trajectory of random walk
We can consider a random walk on a lattice (see a problem at the end of this lecture). Let \( \ell_i \) be the \( i \)th step of the walk. This vector must be one of the bond vectors making the lattice. Starting from the origin, a random walk of \( n \) steps on a lattice would reach
\[
R(n) = \ell_1 + \ell_2 + \cdots + \ell_n. \tag{7.13}
\]
If the lattice spacing is \( a \), then the consideration above (or (7.12)) tells us
\[
\langle R(n)^2 \rangle = na^2. \tag{7.14}
\]

98
We may interpret the trajectory of a random walk as a conformation of a polymer consisting of \(n\) monomers (without any steric interactions among monomer units except perhaps for bond angle constraints). Then, \(R(n)\) is the end-to-end vector of the polymer chain, and the mean square end-to-end distance satisfies (7.14).

7.8 Einstein guessed the Brownian motion is due to thermal motion
That the cause of the Brownian motion is thermal motion of molecules was quantitatively demonstrated for the first time by Einstein in 1905,\(^{124}\) three years before Langevin’s work discussed above. You can understand the original paper in about a month, but not yet, because Einstein invented statistical mechanics by himself and used it to calculate the driving force for Brownian particles.

7.9 Einstein’s approach was consciously mesoscopic
Einstein did not invent the word ‘mesoscopic,’ but his work consciously treated the Brownian particle as a mesoscopic object. This was the reason why his work was not instantly understood as a key paper in thermal physics.

Einstein considered the diffusion process of a collection of Brownian particles. The diffusion flux \(J\) may be written as

\[
J = -D \text{ grad } n, \tag{7.15}
\]
where $n$ is the number density of the Brownian particles. $D$ is defined by this equation (Fick’s law). We computed the diffusion constant in the gas phase though computing the flux.\textsuperscript{125} Einstein did a similar thing for a suspended particle in a fluid medium.

### 7.10 Einstein’s theory of Brownian particle flux

Einstein’s key idea was that a Brownian particle may be treated both as a large molecule and as a tiny macroscopic particle at the same time (i.e., virtually, he introduced the ‘mesoscopic scale’ description of Nature):

(a) Since we regard Brownian particles as molecules, we may apply Dalton’s law of partial pressures. We assume the number density $n$ of the Brownian particles is very small, so they do not interact with each other; we may regard the collection as an ideal gas (the particles are treated microscopically):

$$P = nk_BT.$$  \hfill (7.16)

(This $P$ corresponds to the osmotic pressure due to the solute = Brownian particles as will be discussed in Lecture 20.)

(b) The average of mesoscopic quantities must be understandable macroscopically (i.e., in terms of macroscopic laws); Einstein did not explicitly say this, but as emphasized repeatedly, this is the key mesoscopic feature. Let $f$ be the average force acting on each particle (see Fig. 7.7). The total force acting on the slice in the figure is $nfAdr$, but this must be the same as the force due to the partial pressure difference

$$nfA\, dr = -A[P(r+dr) - P(r)] = -A\, \text{grad} \, P \, dr, \hfill (7.17)$$

![Figure 7.7: The total force acting on the thin slice of thickness $dr = |dr|$ may be understood as the force due to the pressure difference, so $nA\, f\, dr = A(P(r) - P(r + dr)) = -A\, \text{grad} \, P(r)\, dr$, which gives (7.18).](image)

That is,

$$nf = -\text{grad} \, P. \hfill (7.18)$$

\textsuperscript{125}However, do not confuse the formula obtained in 6.11 (i.e., $D = \overline{v}t/3$, which describes the diffusion of gas molecules) and Einstein’s formula for suspended particles in a fluid.
On the average a Brownian particle behaves as a macroscopic particle, so its (average) velocity $v$ due to pushing by $f$ must obey

$$\zeta v = f, \quad (7.19)$$

where $\zeta$ is the friction constant between the particle and the surrounding fluid (drag coefficient) (used in Langevin’s approach 7.4).

The diffusion flux $J$ is

$$J = n v = n f / \zeta \quad (7.20)$$

so (7.18) tells us that (assuming the temperature is uniform)

$$J = -\frac{k_B T}{\zeta} \text{grad } n. \quad (7.21)$$

### 7.11 Einstein’s formula

Comparing (7.21) with the definition of the diffusion constant (7.15), we obtain

$$D = k_B T / \zeta, \quad (7.22)$$

which is called *Einstein’s relation*. This equation allows us to obtain $k_B$ or, since the gas constant $R$ is known, to calculate Avogadro’s constant $N_A$.

Einstein’s original paper used $\zeta = 6 \pi a \eta$, where $a$ is the radius of the Brownian particle, and $\eta$ is the shear viscosity of the fluid. Thus, the original Einstein’s relation reads (often called the *Einstein-Stokes formula*)

$$D = \frac{k_B T}{6 \pi a \eta}. \quad (7.23)$$

Here, Stokes’ law is used that gives the drag force acting on a sphere of radius $a$ moving at velocity $v$ relative to the surrounding fluid: $f = 6 \pi a \eta v$.\textsuperscript{126}

$D \propto k_B T / a \eta$ may be concluded with the aid of dimensional analysis. Since dimensional analysis is quite important, let us derive this relation dimensional-analytically. An introduction to dimensional analysis is available: [http://www.yoono.org/Y_OONO_official_site/LectureSlides_504_files/DAmemo.pdf](http://www.yoono.org/Y_OONO_official_site/LectureSlides_504_files/DAmemo.pdf).

\textsuperscript{126}Its derivation is not very trivial; see, for example, Landau-Lifshitz, *Fluid Dynamics*. 101
7.12 Dimensional analytic ‘derivation’ of Einstein’s formula

The dimension of a quantity $X$ is usually denoted by $[X]$. The basic dimensions are length $L$, mass $M$ and time $T$. For example, $[a] = L$. If you know the unit of a quantity, it is easy (pragmatic) to obtain its dimension from the unit. For example, $J = -D \text{grad} n$. The particle number flux is the number of particles going through a unit area in unit time, so $[J] = 1/L^2 T$, because the number of particles is dimensionless (pure number). $[n] = 1/L^3$. Gradient is essentially differentiation with length, so $[\text{grad}] = 1/L$ (differentiation is something like division).

Therefore, $[J] = 1/L^2 T = [D]/L^4$, so $[D] = L^2 / T$ as we already concluded from the unit of diffusion. We need $[\eta]$. Let us go back to its definition: $J_p = -\eta \text{grad} v$, where $J_p$ is the momentum flux, and $v$ is the velocity. Since the dimension of momentum is $ML/T$, $[J_p] = (ML/T)/L^2 T = M/LT^2$, $[v] = L/T$, so $[\eta] = [J_p]L/[v] = M/LT$. $k_B T$ has the dimension of energy, $[k_B T] = M(L/T)^2$.

Let us determine $D$. In dimensional analysis, first we must itemize all the quantities we believe relevant. In the present example, diffusion should be slow with large $a$ or large $\eta$, and also it is related to thermal motion, so $T$ should matter; $T$ always appears with $k_B$, so we may conclude that $D$ should depend on $a$, $\eta$ and $k_B T$. $[D]$ does not contain $M$, so we should get rid of $M$: $[k_B T/\eta] = (ML^2/T^2)/(M/LT) = L^3/T$. Therefore, $k_B T \eta / a$ must have the same dimension as $D$. Thus, $D \propto k_B T / a \eta$. No other combination is possible.

7.13 Displacement of particles by diffusion

Einstein’s relation (7.22) with (7.9) due to Langevin implies

$$\langle r^2 \rangle = 2D Dt. \tag{7.24}$$

Einstein, before Langevin, derived this equation in a different way (as discussed below), studying the time evolution of the number density $n(r, t)$ of the Brownian particles that obeys the diffusion equation. We know the number density obeys the diffusion equation (6.19):

$$\frac{\partial n}{\partial t} = D \Delta n, \tag{7.25}$$

where $\Delta$ is the Laplacian:

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{7.26}$$

The easiest method to solve (7.25) is to use the Fourier transformation (see Appendix 6A), but here we simply quote the result. If $n$ is normalized with the total number of particles, we get the probability density distribution of the Brownian particles $P(r, t)$. Needless to say, $P$ obeys the same diffusion equation. Let us assume at $t = 0$, all the probability is concentrated at the origin. Then,

$$P(r, t) = \left(\frac{1}{4\pi Dt}\right)^{3/2} e^{-r^2/4Dt}. \tag{7.27}$$
Therefore, after $t$, the mean square displacement must be\footnote{Do not forget that (7.27) is a Gaussian density distribution in $\delta$-space. Since $r^2 = x_1^2 + x_2^2 + \cdots + x_6^2$, (7.27) is actually a product of 6 independent Gaussian density distributions for each orthogonal component $x_i$:}

\[ \langle r(t)^2 \rangle = 2\delta Dt. \] (7.28)

That is, if we observe the mean square displacement of a particle, then the diffusion constant $D$ of the collection of such particles may be measured. Jean Perrin (1870-1942) implemented this measurement and obtained Avogadro’s constant (see a problem at the end of this lecture).\footnote{J. Perrin, *Atoms* (Constable, 1916) translated by D. L. Hammick. Available on line: https://archive.org/details/atomsper00perruoft.}

### 7.14 Einstein’s fundamental idea: summary

Let us summarize Einstein’s fundamental idea, which is the key idea of the current nonequilibrium statistical mechanics:

Microscopic fluctuations can build up mesoscopic fluctuations whose dynamics is on the average governed by the laws of macroscopic time evolution.

The Brownian motion is a mesoscopic motion that is a result of building up of microscopic fluctuations. Its decay is described by a macroscopic dissipative dynamics.

This was later more clearly stated by Onsager as the *regression hypothesis*.

### 7.15 Noise in the Langevin approach

We started with Langevin’s theory based on a stochastic differential equation. If there is an external force $F$, it may be generalized as

\[ m\frac{d^2r}{dt^2} = -\zeta \frac{dr}{dt} + F + w. \] (7.29)

Usually, we may assume that $m$ is small and $\zeta$ is large (i.e., the over-damped condition). Then, we may ignore the acceleration term, and the equation is rewritten

\[ \langle x_i^2 \rangle = \frac{2}{d} \frac{8}{6} \pi D t. \]

This is a one-dimensional Gaussian density distribution, so we immediately see $\langle x_i^2 \rangle = 2Dt$. Consequently, we have $\langle r^2 \rangle = 6\langle x_i^2 \rangle = 2\delta Dt$.\footnote{\[ 128 J. Perrin, *Atoms* (Constable, 1916) translated by D. L. Hammick. Available on line: https://archive.org/details/atomsper00perruoft.}
as
\[
\frac{dr}{dt} = \frac{F}{\zeta} + \nu, \quad (7.30)
\]
where \(\nu\) is a noise. This equation is also called the Langevin equation (actually, this is the usual one).

Now, we wish to model the noise \(\nu\). We assume its all components are statistically independent, so let us study its \(x\)-component \(\nu_x\) as a representative. \(\nu_x(t)\) as a function of time changes quite rapidly and erratically, so we assume its ensemble average to satisfy
\[
\langle \nu_x(t) \rangle = 0, \quad (7.31)
\]
and
\[
\langle \nu_x(t)\nu_y(s) \rangle = A\delta(t - s), \quad (7.32)
\]
where \(A\) is a positive numerical constant. This implies that there is no memory in noise (different times are uncorrelated). Since three components are uncorrelated,
\[
\langle \nu_x(t)\nu_y(s) \rangle = 0, \text{ etc.} \quad (7.33)
\]

In order to determine \(A\), let us solve (7.30), assuming that the particle starts from the origin:
\[
r(t) = \int_0^t ds \, \nu(s). \quad (7.34)
\]
From this we obtain
\[
\langle r^2(t) \rangle = \int_0^t ds \int_0^t ds' \langle \nu(s) \cdot \nu(s') \rangle = 6A \int_0^t ds \int_0^t ds' \delta(s - s') = 6A \int_0^t ds = 6At. \quad (7.35)
\]
Comparing this with (7.9), we can determine \(A\) as
\[
A = \frac{2k_B T}{\zeta} = 2D. \quad (7.36)
\]
This is called a fluctuation-dissipation relation, determining the noise (i.e., fluctuation) ‘amplitude (squared)’ \(A\) in terms of temperature \(T\) and friction constant \(\zeta\) (i.e., dissipation).
7.16 Qualitative understanding of fluctuation-dissipation relation

Qualitatively, the fluctuation-dissipation relation (FDR) may be understood as follows (see Fig. 7.8). If the force in (7.30) is conservative (has a potential $U$) as

$$ F = -\text{grad} U $$

(7.38)

our Langevin equation becomes

$$ \frac{dr}{dt} = -\frac{\nabla U}{\zeta} + w. $$

(7.39)

Since the particle tend to be trapped in $U$, in equilibrium, we should expect the Boltzmann distribution $\propto e^{-\beta U}$ for the particle position $r$ when the system described by (7.29) reaches an equilibrium state.

![Figure 7.8: Illustration of the fluctuation-dissipation relation for a Brownian particle in a potential $U$. $P$ is the probability density to find the particle at the given location with: Blue: too small noise; Green: just right noise; Red: too large noise for a given $T$ and $\zeta$. To reproduce the correct equilibrium state, the noise must be carefully chosen.](image)

For example, if the viscosity of the suspending liquid is large, $\zeta$ is large, and as can be seen from (7.39), the effect of the systematic force due to the potential energy becomes relatively small. If the noise amplitude is not reduced appropriately, then obviously the distribution would spread too much. In other words, when the ambient liquid is viscous, large noise pushes the particle away from the potential minimum. Before the particle reaches the minimum, another noise kicks the particle further away from the potential minimum. Thus, a larger damping effect must be balanced with a smaller noise amplitude for the Langevin equation to describe the equilibrium state correctly. This is actually (5) in the summary of observations of the Brownian motion at the beginning of this lecture.

7.17 Smoluchowski equation and Einstein relation

The fluctuation-dissipation relation is usually derived from the condition that the
equilibrium distribution of the particles is correctly described by the Boltzmann factor.\textsuperscript{129} To close our discussion of the Brownian motion the approach to the fluctuation-dissipation relation through the front door will be explained.

Let us derive the transport equation for the number density. This is what Einstein did when there was no force. Now with an external force, how should we proceed? Since our system is linear, the force is not strong. Thus the flux must be the superposition of the flux due to the gradient of $n$ and the one driven by the force. The latter reads (recall the flow velocity $\times$ the density of the quantity transported is the flux, and the flow velocity due to the driving by the force due to the potential $U$ reads $\zeta v = -\nabla U$)

$$n \times v = -n\frac{1}{\zeta}\nabla U.$$  
\text{(7.40)}

Therefore, the total flux reads with the aid of (7.37)

$$\mathbf{J} = -\frac{A}{2}\nabla n - \frac{1}{\zeta}n\nabla U.$$  
\text{(7.41)}

The conservation of particles reads

$$\frac{\partial n}{\partial t} = -\text{div} \mathbf{J}.$$  
\text{(7.42)}

\subsection*{7.18 Smoluchowski equation and fluctuation-dissipation relation}
Combining (7.42) and the formula for the flux (7.41) gives the following equation called the \textit{Smoluchowski equation}

$$\frac{\partial n}{\partial t} = \nabla \cdot \left(n\frac{1}{\zeta}\nabla U + \frac{A}{2}\nabla n\right).$$  
\text{(7.43)}

In equilibrium, the time derivative must vanish, so

$$n\frac{1}{\zeta}\nabla U + \frac{A}{2}\nabla n = 0,$$  
\text{(7.44)}

because this must vanish far away from the potential, or

$$\frac{2}{A\zeta}\nabla U + \nabla \log n = 0.$$  
\text{(7.45)}

\textsuperscript{129}Notice that even the above ‘back door’ derivation imports the equilibrium result as the equipartition of energy or the formula for the pressure.
This implies that

\[ n \propto e^{-2U/A\zeta} \]  

(7.46)

which must be proportional to the Boltzmann factor 5.11. Consequently, we must conclude that

\[ A\zeta/2 = k_B T. \]  

(7.47)

\[ A = 2k_B T/\zeta \]

is the fluctuation-dissipation relation, but it is the Einstein relation itself.
Appendix 7A: How to obtain (7.27)

Use of Fourier transformation is the best. Fourier transformation $\mathcal{F}$ is defined as follows (in 3-space):

\[ [\mathcal{F}f](k) \equiv ˜f(k) = \left(\frac{1}{2\pi}\right)^3 \int d^3x \, f(x) \, e^{ik \cdot x}. \tag{7.48} \]

Notice that differentiation becomes multiplication:

\[ [\mathcal{F}(\nabla f)](k) = -i k ˜f(k). \tag{7.49} \]

This can be demonstrated by (essentially an integration by parts)

\[ [\mathcal{F}(\nabla f)](k) = \left(\frac{1}{2\pi}\right)^3 \int d^3x \, \nabla f(x) \, e^{ik \cdot x} \]

\[ = \left(\frac{1}{2\pi}\right)^3 \int d^3x \left[ \nabla \left( f(x) e^{ik \cdot x} \right) - i k f(x) e^{ik \cdot x} \right]. \tag{7.50} \]

The first term in the second line above vanishes (assuming $f$ vanishes at infinity), and we get (7.49).

From $f$ we can recover $g$ by the inverse transformation:

\[ f(x) = [\mathcal{F}^{-1} ˜f](x) = \int d^3k \, ˜f(k) \, e^{-ik \cdot x}. \tag{7.51} \]

Let us Fourier transform the diffusion equation (7.25). We get

\[ \frac{d\tilde{n}(t, k)}{dt} = -D k^2 \tilde{n}(t, k). \tag{7.52} \]

This is an ordinary differential equation ($k$ is a mere parameter). The initial condition is $n(0, x) = δ(x)$ (i.e., initially all the particles are at the origin). Its Fourier transform is

\[ \tilde{n}(0, k) = 1/8\pi^3. \tag{7.53} \]

Therefore, the solution to (7.52) is

\[ \tilde{n}(t, k) = \frac{1}{8\pi^3} e^{-Dk^2t}. \tag{7.54} \]

Now, we inverse-transform this to get

\[ n(t, r) = \int_{\mathbb{R}^3} d^3k \, \frac{1}{8\pi^3} e^{-Dk^2t - ik \cdot r} \]

\[ = \frac{1}{8\pi^3} \int_{\mathbb{R}^3} d^3k \, e^{-tD(k+ir/2Dt)^2-r^2/4Dt} \]

\[ = \frac{1}{8\pi^3} \left(\frac{\pi}{tD}\right)^{3/2} e^{-r^2/4Dt} = \left(\frac{1}{4\pi tD}\right)^{3/2} e^{-r^2/4Dt}. \tag{7.57} \]
The procedure from (7.55) to (7.56) is the completion of square in the exponent (see 5.8), and then the calculation from (7.56) to (7.57)
Q7-1. One experiment replicating Perrin’s experiment in a modern setting uses polystyrene particles of diameter (i.e., $2a$) 0.5 $\mu$m suspended in a buffer solution of viscosity $\eta = 1.03 \times 10^{-3}$ Pa-s at $T = 300$ K. A two-dimensional stage was recorded by a microscope with a CCD camera, and its $x$ coordinate is measured as a function of time. The mean square average displacement in $x$ is observed as $\langle x^2 \rangle = 15.6 \times 10^{-13}t$ m² after $t$ seconds. Assuming that you know the gas constant $R = 8.31$ J/mol·K, estimate Avogadro’s constant $N_A$.

Soln.
The relation we use is $\langle x^2 \rangle = 2D_t$ and the Einstein-Stokes relation $D = k_B T / 6\pi a \eta$. Therefore, $k_B = (3\pi a \eta / T) \times 15.6 \times 10^{-13} = 1.26 \times 10^{-23}$, or $N_A = R / k_B = 6.58 \times 10^{23}$.

Q7-2. On a triangular lattice or a honeycomb lattice (see Fig. 7.9) with the same edge (i.e., lattice bond) length $\ell$ is a random walker.

The walker starts from the origin $O$ and walks along the edges. At every second she chooses randomly any of the edges connected to her current position and moves to the nearest neighbor lattice point along the chosen edge. You can assume that she completely forgets at what lattice point she was previously (i.e., all the steps are statistically independent). If her $i$th step displacement is denoted by vector $a_i$, the total displacement during time $t$ seconds is given by

$$R = a_1 + a_2 + \cdots + a_t.$$  \hfill (7.58)

Here, all the step vectors $a_i$ are lattice bond vectors.

(1) After $t$ seconds on which lattice ($T =$ triangular or $H =$ honeycomb) can she be further away from the origin on the average? That is, choose the correct relation from the following and justify your choice: $T > H$, $H > T$ or $H = T$.

(2) Now, on the triangular lattice due to a strong wind blowing constantly in the $+x$ direction, the walker tends to choose $+x$ direction with probability 0.5, but still chooses the remaining five directions randomly (with probability 0.1 for each).

(a) What is the average position ($x$ and $y$ coordinates) of the walker after $t$ seconds?

(b) What is the variance of the $y$-coordinate after $t$ seconds?
(c) What is the mean square displacement $\langle R^2 \rangle$ of the walker after $t$ seconds?

**Soln.**

(1) Thanks to the statistical independence of steps and the average step displacement being zero (i.e., $\langle a_i \rangle = 0$, so $\langle a_i a_j \rangle = \ell^2 \delta_{ij}$), we obtain

$$\langle R^2 \rangle = \sum_{i=1}^{t} \langle a_i^2 \rangle.$$ 

Here, $\langle \rangle$ is an ensemble average. Obviously, $\langle a_i^2 \rangle = \ell^2$, so $\langle R^2 \rangle = t\ell^2$. Does this calculation depend on spatial dimension or the lattice structure?

We have $H = T$.

This might be slightly counterintuitive, because the honeycomb lattice walk seems less ‘zig-zag’ than the other case. Do not forget that there is a significant probability $(1/3)$ to retrace the immediate-past step to return to the same position the walker was 2 sec ago.

(2) (a) The position after $t$ seconds is given by (7.58). Therefore, the average position is $\langle R \rangle = t \langle a_1 \rangle$.

$$\langle a_i \rangle = 0.5(\ell, 0) + 0.1\ell \sum_{k=1}^{5} \left( \cos \frac{k\pi}{3}, \sin \frac{k\pi}{3} \right),$$

but from the symmetry without actual calculation

$$\langle a_i \rangle = 0.5(\ell, 0) + 0.1(-\ell, 0) = (0.4\ell, 0).$$

Therefore, $\langle R \rangle = (0.4\ell t, 0)$.

(b) Let us write $R = (X, Y)$. Then, $Y = \sum_{i=1}^{t} y_i$, where $y_i$ is the $y$-component of the $i$th step vector. We know $\langle Y \rangle = 0$, so using the statistical independence of steps, we have

$$V(Y) = \langle Y^2 \rangle = \sum_{i=1}^{t} \langle y_i^2 \rangle,$$

where

$$\langle y_i^2 \rangle = 0.6 \times 0 + 0.4 \left( \ell \sin \frac{\pi}{3} \right)^2 = 0.3\ell^2$$

Therefore, $V(Y) = 0.3\ell^2 t$.

(c) We need

$$\langle R^2 \rangle = \sum_{i=1}^{t} \langle a_i^2 \rangle + \sum_{i \neq j} \langle a_i \cdot a_j \rangle.$$ 

Although each step is statistically independent (so you may write $\langle a_i \cdot a_j \rangle = \langle a_i \rangle \cdot \langle a_j \rangle$), its average is not zero in this case, so you cannot ignore the cross terms. There
are $t(t - 1)$ cross terms, but they are all the same:

$$
\langle a_i \cdot a_j \rangle = \langle a_i \rangle \cdot \langle a_j \rangle = \langle a_1 \rangle^2.
$$  \hfill (7.59)

We have already computed $\langle a_1 \rangle = (0.4\ell, 0)$. Obviously, $\langle a_i^2 \rangle = \ell^2$

$$
\langle R^2 \rangle = \ell^2 t + 0.16\ell^2 t(t - 1).
$$  \hfill (7.60)

What is the variance of $\mathbf{R}$?
Q7-3. Suppose two identical Brownian particles are released from the same point on a two dimensional stage. Assume that the diffusion constant of the particles is $D = 1.5 \times 10^{-12}$ m$^2$/s.

What is the root mean-square displacement of the average position of these two particles after 1 hr.

**Soln.**

$$
(1/4)\langle (\mathbf{r}_1 + \mathbf{r}_2)^2 \rangle = (1/2)\langle \mathbf{r}_1^2 \rangle = dDt = 2 \times 1.5 \times 10^{-12} \times 3600 = 1.08 \times 10^{-8}.
$$

Therefore, the root mean-square displacement of the average position is $1.04 \times 10^{-4}$ m $\approx 104 \, \mu$m.

To answer such problems, first itemize what you need: two Brownian particles are mentioned, so we need two position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Since the position of the center of mass is asked, let us express it as $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$. Then, do simple algebra without thinking:

$$
\langle \mathbf{R}^2 \rangle = \frac{1}{4}\langle \mathbf{r}_1^2 + \mathbf{r}_2^2 + 2\mathbf{r}_1 \cdot \mathbf{r}_2 \rangle = \frac{1}{4}(\langle \mathbf{r}_1^2 \rangle + \langle \mathbf{r}_2^2 \rangle + \langle 2\mathbf{r}_1 \cdot \mathbf{r}_2 \rangle).
$$

Now, you must look at the result and think about a bit of the actual situation. Both the particles are identical, so the average should not depend on particles $\langle \mathbf{r}_1^2 \rangle = \langle \mathbf{r}_2^2 \rangle$, and as already noted far before Einstein, two Brownian particles are statistically independent if apart more than their sizes. We study a long time behavior, so, except for a very short time near the starting point, these two particles are statistically independent. Consequently,

$$
\langle \mathbf{r}_1 \cdot \mathbf{r}_2 \rangle = \langle \mathbf{r}_1 \rangle \cdot \langle \mathbf{r}_2 \rangle.
$$

The space is isotropic, so there is no preferred direction to wander: $\langle \mathbf{r}_1 \rangle = \langle \mathbf{r}_2 \rangle = 0$. Hence,

$$
\langle \mathbf{R}^2 \rangle = \frac{1}{2}\langle \mathbf{r}_1^2 \rangle.
$$

The rest is as above.

Q7-4. There are two proteins of mass $m$ and $M$. Let us assume that the protein molecules are spherical and its average densities are the same. We know $M/m = 100$.

For a smaller molecule to diffuse across a fixed length $L$ in a cell it takes 0.23 s on the average. What is the best guess of the time needed for the larger protein to diffuse across the same distance $L$?

**Soln.**
\[ \langle r^2 \rangle = 2D t, \quad D = k_B T / 6 \pi a \eta. \]  This means \( D t \) is the same, so \( t/a \) is the constant. Since we assume that the proteins are spherical and with the same density, \( a \propto M^{1/3} \). That is, \( t/M^{1/3} \) is constant. Hence, \( t = 0.23(M/m)^{1/3} = 1.07 \) s.

As warned repeatedly, do not use the gas phase formulas to calculate the diffusion constant in liquids.
8 Macrosystems

Summary
* If numerous particles gather, important observables are extensive (additive) or intensive. Also even though underlying mechanics is reversible, macroscopic systems exhibit irreversibility.
* For a macroscopic system its total mechanical energy is additive with high precision.

Key words
Internal energy, reversibility of mechanics, Poincaré recurrence

What you should be able to do
* Explain why the total mechanical energy of a macroscopic system is additive.
* Explain why irreversibility naturally occurs in systems with many particles.

8.1 How to describe a macroscopic system in mechanics
We do not need any special way to describe a macroscopic system, if we wish to describe it purely mechanically. Mechanical entities are atoms and molecules, so a system is mechanically described by the system Hamiltonian whose independent variables are position and momentum vectors of particles.

The Hamiltonian of a system consisting of \( N \) point particles of mass \( m \) interacting with a potential energy \( U(x_1, \ldots, x_N) \) has the following form

\[
H = \sum \frac{1}{2} m \dot{x}_i^2 + U(x_1, \ldots, x_N).
\] (8.1)

The first terms describe the kinetic energy \( K \) (here, \( K \) is the total kinetic energy). Usually, we may assume that \( U \) depends on the mutual positions of the particles and not on the absolute positions of the particles. The value of \( H \) is the total mechanical energy of the system.

Since we are interested in the ‘intrinsic’ properties of the system, we are not interested in the overall translation and rotation. Thus, we are interested in the Hamiltonian of the system observed from the coordinate system relative to which the system does not exhibit any overall translational and rotational motion (the co-moving coordinate system). The total mechanical energy of the system observed by the co-moving observer is understood as the ‘intrinsic’ mechanical energy of the system.
8.2 Conservation of mechanical energy and the first law of thermodynamics

Unless there is an exchange of energy with the external world, the total ‘intrinsic’ mechanical energy (8.1) of a system sitting still relative to the observer should obviously be conserved according to the conservation of mechanical energy. In thermodynamics the total ‘intrinsic’ mechanical energy is called the internal energy. Thus, the internal energy of a system must be a conserved quantity. This is the essence of the first law of thermodynamics. This was recognized by Carnot, Mayer, Helmholtz and others, but Helmholtz most clearly recognized the first law as a consequence of the conservation of mechanical energy, especially due to the fact that intermolecular interactions have potential functions.

8.3 Two crucial features of macroscopic systems

What are the salient features of a system consisting of numerous particles? Two features come to our mind: additivity of energy and irreversibility of time evolution.

8.4 Additivity of energy

The usual intermolecular interaction decays spatially sufficiently quickly (Fig. 8.1).

Figure 8.1: The intermolecular force potential. The repulsive portion is very steep (any steep function will do to describe it, say, $1/r^{12}$), which is due to electron-cloud overlap. The attractive portion is $1/r^6$, which is due to the induced dipole-dipole interaction (the London force). Roughly speaking, the binary intermolecular force is characterized by the repulsive (or hardcore) diameter $d$ (the representative length scale) and the depth of the potential well $\varepsilon$ (the representative energy scale).

---

$^{130}$If there are electromagnetic effects, this energy must be expanded to include the electromagnetic energy.

$^{131}$Strictly speaking, thermodynamics discusses the systems in equilibrium, so the first law is a restricted version of the law of conservation of energy.

$^{132}$Since no one can verify all the particles indeed obey microscopic mechanics, a more precise statement is that the empirically established first law strongly suggests that the microscopic mechanical model of a system is in terms of conserved intermolecular forces.
If a system volume is split into two $V_1 + V_2$ with a simple boundary surface,\textsuperscript{133} the sum of the total mechanical energies of the volumes is very close to the total mechanical energy of the whole system before splitting.

If the interaction energy between two particles decays faster than $r^{-d}$ in $\delta$-space, where $r$ is the interparticle distance,\textsuperscript{134} then the total interaction energy of a single molecule $i$ near the boundary of $V_1$ with those in $V_2$ may be estimated as (see Fig. 8.2)

\[
\sum_{j \in V_2^*} \frac{1}{r_{ij}^{d+\varepsilon}} \approx \int_{V_2^*} dy \frac{1}{|y|^{d+\varepsilon}} \propto \int_{\delta}^{L} R^{d-1} dR \frac{1}{R^{d+\varepsilon}} \sim \int_{\delta}^{L} dR \frac{1}{R^{1+\varepsilon}} \sim L^{-\varepsilon} + \text{const.} \tag{8.2}
\]

Here, $r_{ij}$ is the distance between particles $i$ and $j$, $V_2^*$ is the subset of $V_2$ such that the portion within distance $\delta$ from particle $i$ is removed from $V_2$. This calculation shows that the relative contribution of the interaction energy that depends on the system size $L$ becomes smaller as the system size becomes bigger. Of course, we cannot ignore the close-range contributions from the portion within distance $\delta$ of the particle as well as the constant term in (8.2), but even if we collect such contributions from all the molecules on or near the boundary, they are only proportional to $L^{d-1}$ (i.e., the surface area), so we may ignore them relative to the bulk energy $\propto L^d$ for macroscopic systems.

Thus, if the interaction potential decays faster than $r^{-d}$, then we may assume that the total energy is proportional to the volume of the system.

The Coulomb and gravitational interaction energies decay as $1/r$. For the Coulomb interaction, thanks to the shielding effect, if the system is charge-neutral, the effective interaction decays exponentially, so we need not worry about it. For the gravity, there is no way to shield it, but the usual macroscopic objects we are interested in

\textsuperscript{133}not a fractal surface, for example; we say we split the volume into two volumes in a \textit{van Hove} way.

\textsuperscript{134}Since the interactions are not totally binary, precisely speaking, we need to assume that the total interaction energies among particles is about the same order as the total contribution of the binary interactions. This is a bit delicate, but usually this is true.

117
in statistical physics is not huge (usually about 1 m\(^3\) or less with not a huge density like a neutron star), so we may ignore it (unless we discuss sedimentation equilibrium; even in such cases we may safely ignore gravitational interactions among the particles in the system).\(^{135}\)

**8.5 Why we are interested in additive quantities**

To study a macroscopic system we often study additive observables that is proportional to the system size (think of internal energy). This is natural, because such observables are big for big systems. If an observable becomes smaller for larger systems, we need not pay attention to such quantities to understand the macroscopic world. Thus, we are interested in observables that are independent of the system size (intensive quantities) and those proportional to the amount of materials in the system (extensive quantities).

**8.6 Time-reversal symmetry of mechanics**

The world of mechanics is time-reversal symmetric (the world in which the movies played backward do not look funny). In the case of classical mechanics, Newton’s equation of motion of a closed (isolated) system is an autonomous differential equation of second order without any first order derivative: For an \(N\)-particle system, generally we have

\[
\frac{d^2 x_i}{dt^2} = F_i(x_1, \cdots, x_N) \tag{8.3}
\]

for \(i = 1, \cdots, N\). Since the forces \(F_i\) are \(t\) independent in the closed system, \(t \to -t\) does not change the equations.

The Schrödinger equation for an isolated system reads

\[
i\hbar \frac{\partial \psi}{\partial t} = H\psi, \tag{8.4}\]

where \(H\) is a Hermitian operator independent of time. In this case \(t \to -t\) might seem to alter physics, but what we observe is real, so the physics must be intact under complex conjugation.\(^{136}\) Thus, quantum physics is also intact under time reversal.

**8.7 In the long run we are all dead**

But in the long run we are all dead and will never be resurrected. The world we live

---

\(^{135}\)If the system really becomes huge, our ordinary statistical thermodynamics does not work.

\(^{136}\)Hermitian conjugation, more precisely.
8.8 Irreversibility from mechanics?
All the ambitious young men (Boltzmann 1844-1906, Einstein, 1879-1955, ...) tried to explain irreversibility from mechanics.

Boltzmann derived in 1872\textsuperscript{137} from a pure microscopic mechanical description of atoms an equation (called the Boltzmann equation) that governs the irreversible time evolution of dilute gasses by ignoring some statistical correlations. His colleague Loschmidt (1821-1895) asked why Boltzmann could derive an irreversible equation from a reversible equation. This question made Boltzmann realize that his derivation included a sort of coarse-graining. Boltzmann explained that the initial ‘non-equilibrium’ states always contain more order (so inevitably subtle correlations as well), so the time evolution always drives the system to the less ordered direction; thus his equation correctly captures this tendency.

Then, in 1896, came Zermelo (1871-1953), an assistant to Planck (1858-1947) those days, who, utilizing Poincaré’s recursion theorem,\textsuperscript{138} pointed out that Boltzmann’s argument was logically flawed: even if the system is coarse-grained, sooner or later the state of a closed system returns to a state indefinitely close to the starting state, so no irreversibility occurs. Boltzmann admitted the flaw, but since he was a theoretical physicist, he responded: “Young man, you know math, but you don’t know physics. Think how long it takes for that to happen? It would take far longer than the age of the universe even for a very small system.”

8.9 What is the lesson?
What we have learned from these debates are:
(1) Very often the initial state is special (away from equilibrium) so (even following the pure mechanics) for a long time irreversible behaviors are observed. The reason why the initial state cannot be recovered in various theories is that they discard subtle correlations (coarse-grain the system).
(2) However, if we can wait for long enough, any finite system (even after coarse-graining) almost comes back to its initial special state, but the required time is enormous, and we never experience this for macroscopic objects.

\textsuperscript{137}[1872: Yellowstone NP established as the first NP in the world; G. Elliot: Middlemarch (in which Brown’s famous booklet on Brownian motion showed up); C Monet: Impression Sunrise]

\textsuperscript{138}Irrespective of time reversibility or irreversibility, a measure-theoretical dynamical system can return to a state indefinitely close to its initial condition.
8.10 A toy model illustrating the lessons

A toy model can illustrate these points. Suppose a point is going around a unit circle with period 1 uniformly. The point is subjected to a noise that makes its angular speed fluctuate (see Fig. 8.3).

![Figure 8.3: Ensemble of points with angular speeds slightly fluctuating around $2\pi$ rad/s. The averaged position spirals toward the center.](image)

If there are only two such oscillators, their average position may become close to the origin, but then the average recovers its original amplitude fairly easily. If we have many ($N \gg 1$) such oscillators, after the average becomes close to zero, it stays small for an enormously long time, and then will return close to the original value. The waiting time for this recovery is likely to be of order $e^{cN}$, where $c$ is a positive constant of order 1. [I do not know the precise quantitative results.]

Thus, as long as the system is finite, Zermelo is right, but as to the waiting time Boltzmann is right.
9 Thermodynamics: Principles

Summary
* Phenomenological approach is a respectable and basic method for understanding the world
* Equilibrium states may be described by the thermodynamic coordinates consisting of the internal energy and the work coordinates. The space spanned by the thermodynamic coordinates is called the thermodynamic space.
* Even if the change from one equilibrium state to another is irreversible, by devising a quasistatic path between them, changes in thermodynamic quantities may be computed thermodynamically.
* The first law is essentially the conservation of energy, but to describe it in terms of small number of macroscopic variables, changes (or processes) must be sufficiently slow.
* The second law implies that the thermodynamic space is foliated into adiabats = constant-entropy surfaces.

Key words
Phenomenological approach, zeroth law, fourth law, thermodynamic coordinates, work coordinates, thermodynamic space, quasiequilibrium process, reversible process, state function, thermal contact, thermal equilibrium, temperature, conjugate pair, Clausius’ law, Kelvin’s law, Planck’s law, adiabatic process, adiabat, Gibbs relation

What you should be able to do
* Explain why the thermodynamic coordinates are privileged coordinates.
* Understand that all three expressions of the second law mentioned here are equivalent.
* Demonstrate that we can introduce a state function called entropy. Also be able to explain its relation to heat.

9.1 Mechanics is not enough
As we have discussed in the last lecture, important characterization of macroscopic systems, extensivity of internal energy and irreversibility, may be understood from mechanics, so perhaps you may think mechanics can explain everything thermal as

139There are many for the present section, but they are important, so you should try to memorize them at least once.
well. Unfortunately, however, we cannot confine ourselves to the discussion of closed (or isolated) systems. For example, we must discuss heat transfer, which is hard to describe in terms of mechanics.\footnote{Also we saw that the effect of the external world of the system cannot be completely eliminated. The long time limit ($t \to \infty$) and the ‘external noise zero’ limit (i.e., the pure mechanical limit) are not commutative, so purely mechanical description may well fail to describe the system reaching equilibrium after a long time.}

We have learned that, although Brownian dynamics may well be due to underlying particle mechanics, we could set up a reasonable model (Langevin’s equation) with the aid of the fluctuation-dissipation relation without directly referring to the underlying mechanics. This suggests that with a small number of postulates we can describe the key features of the macroscopic level. Such a description is a phenomenological description (phenomenology).

### 9.2 What is phenomenology

A phenomenological description of macroscopic systems is a description solely in terms of quantities that may be observed, described and defined on the macroscopic scale. If we could make a closed (complete) system of theory in terms of such quantities, the result is called a \textit{phenomenology}. The phenomenology of macroscopic systems in equilibrium we now have is called thermodynamics. The reader must clearly recognize that, in contrast to (the supposedly more fundamental) statistical mechanics, thermodynamics survived the quantum revolution unscathed; actually it helped launching the revolution. Quantum mechanics has no problem with thermodynamics (for now), but even if quantum mechanics will be replaced by something else in the future, thermodynamics will remain intact.

In physics, ‘phenomenology’ has not necessarily been respected; it is almost a pejorative (e.g., in high-energy physics). However, notice that when underlying microscopic descriptions are impossible or only approximate, phenomenology may be the only realistic rational means for the human-beings to understand the world. A phenomenology is not an approximate way to understand the world nor a crude version of something more accurate; thermodynamics is not an approximation of a certain more advanced and accurate theory.

### 9.3 Equilibrium states

A macroscopic system (a system with extremely many particles\footnote{From the strictly macroscopic phenomenological point of view, since whether atoms and molecules exist or not is an irrelevant issue, it is more logical to say, “a}) can take a spe-
cial state called a *thermodynamic equilibrium state* (called an equilibrium state for simplicity). To specify the state, we introduce the concept, *generalized isolation*: a system is subjected under a generalized isolation condition if it is isolated but may be coupled to uniform field(s) conjugate to work coordinates (to be defined shortly). A macrosystem is in an *equilibrium state*, if it is left under a generalized isolated condition and eventually reaches a time-independent state. Macro-observables observed at any time in equilibrium will take unchanging values forever.

A macroscopic system in equilibrium is partitioning-rejoining invariant: if a macroscopic system in equilibrium state is divided into two halves (of about the same sizes), the halves are themselves in equilibrium and if they are joined again, the result is indistinguishable from the original system as long as the thermodynamic observables are concerned (Fig. 9.1).

![Figure 9.1: Partitioning-rejoining invariance of equilibrium states](image)

**Figure 9.1:** Partitioning-rejoining invariance of equilibrium states

Most textbooks are wrong: A traditional characterization of an equilibrium state is as follows: if a macrosystem is isolated and is left undisturbed for a long time, it would reach a macroscopic state (which is characterized by macroscopic observables) which would not change any further (if observed through macroscopic observables). This final state is a thermodynamic equilibrium state.

Unfortunately, many states thermodynamics wish to consider cannot be reached around us that we can see by our naked eyes” (as explained in Section 1, this characterization is actually scientific), but the book should be practical as well, so anything useful will be used to understand thermal physics.

142 **(Equilibrium: another possible characterization)** Equivalently, a macrosystem is in an equilibrium state, if we can devise a (macroscopically) constant and spatially uniform environment (without any dissipative currents) into which we can embed the system (with appropriate boundary conditions) and still the state does not change in time. Total isolation is a possible environment. However, not all the equilibrium states of a given system may be realized under the total isolation condition. Notice that this characterization of equilibrium never requires the isolation of a system from the external world. Also it never asks how the equilibrium state is realized. The state may be prepared in contact with a heat bath, for example.
this way. Thus, the traditional definition is, if not wrong, grossly incomplete.

9.4 The fourth law of thermodynamics
As already discussed intuitively, macroscopically important observables are extensive or intensive. All the thermodynamic observables are either extensive or intensive. This is called the fourth law of thermodynamics.

Notice that partitioning-rejoining invariance in 9.3 makes the fourth law operationally meaningful.

9.5 Thermodynamic coordinates, a privileged set of variables
Since thermodynamics must be a phenomenology on the macroscopic scale, to construct a closed (self-contained) theoretical framework, we must carefully choose the physical quantities we deal with. The most fundamental of them are the thermodynamic coordinates. They are extensive quantities (= additive quantities) and consist of internal energy $E$ and other variables (called work coordinates) that we can observe and control mechanically (electrodynamically) macroscopically. The system volume $V$ is often among them. For a magnetic system, magnetization $M$ is included.

Thermodynamic coordinates are a very special set of variables to describe equilibrium states that is privileged in the following sense:
(i) To understand thermodynamic coordinates we need only (macroscopic) mechanics and electromagnetism. The work that a system does or that a system is done can be described through the changes of these coordinates and is quantified solely electrodynamically. Notice that we need not understand what ‘heat’ or ‘temperature’ is. Thus, as long as the thermodynamic properties are concerned, a macroscopic system is regarded as a black box with mechanically controllable handles.
(ii) The thermodynamic coordinates uniquely specify the equilibrium state. You will understand the meaning of this statement from Figure 9.2. Notice that $T$ or $P$ is not included in the thermodynamic coordinates.

The space spanned by the thermodynamic coordinates is called the thermodynamic space. For a given macroscopic system, its each equilibrium state uniquely corresponds to a point in the thermodynamic space of its own.\textsuperscript{143}

\textsuperscript{143}Some readers might question that there are much more macroscopic observables we can observe for a given object, shapes, orientation, etc. Precisely speaking, thermodynamic states are equivalence classes of macroscopically distinguishable states according to the values of the thermodynamic coordinates.
Figure 9.2: A-C contain the same amount of water at 0°C under 1 atmospheric pressure. However, their internal energies are distinct; A has the least internal energy and C the most. In elementary thermodynamics, often the temperature \( T \) appears as a key variable instead of internal energy \( E \), but these examples clearly tell us that \( T \) cannot distinguish equilibrium states that are distinct. Analogously, in the liquid-gas phase transition under constant pressure \( P \) and \( T \), \( E \) and \( V \) change. These examples indicate that thermodynamic coordinates are the fundamental ‘privileged’ set of variables to describe thermodynamic equilibrium state; generally speaking intensive variables such as \( T \) and \( P \) fail to describe states uniquely.

9.6 Quasistatic processes

Thermodynamics wishes to study changes of equilibrium states by various processes. Not all (actually most) processes allowed to the system cannot be described in the thermodynamic space, because every point actually realizable in the thermodynamic space describes an equilibrium state of the system. Only processes that are extremely (infinitesimally) close (experimentally indistinguishably close) to equilibrium states at every moment may be expressed as continuous curves in the thermodynamic space (Fig. 9.3). In order for a process to be always close to equilibrium states it must be sufficiently slow.

Figure 9.3: A and B are equilibrium states. A quasistatic process connecting A and B is in the thermodynamic space. From A to B a process need not be quasistatic. Then, such a process cannot be described in the thermodynamic space (red).

However, the slowness of a process does NOT guarantee reversible changes.\(^\text{144}\) A

\(^{144}\)Think of a hot coffee in a thermos.
**quasistatic processes** is a process along which both the system and its environment are (infinitesimally) close to equilibrium and can retrace their evolution precisely, i.e., ‘reverse their footsteps.’ Thus, a quasistatic process is also called a **retraceable processes**. Here, ‘retracing’ means that, after retracing the process, the world returns exactly to the original (macro)state. Thus, ‘retraceable’ means that, after retracing, the world returns to the state before the process occurs.

### 9.7 State functions
If the value of a macroscopic quantity of an equilibrium state is uniquely specified by the corresponding point in the thermodynamic space, the macroscopic quantity is called a **state function**. That is, any observable that is a function defined on the thermodynamic space is a state function. Its value is indifferent to how the state is realized. For example, the equilibrium volume of a system is a state function; temperature is another example.

### 9.8 Simple system
An equilibrium system need not be spatially homogeneous at the macroscopic scale. If a system is spatially homogeneous, we call it a **simple system**.

### 9.9 Compound system
Two or more simple systems considered as a single system with or without certain interactions among them is called a **compound system**. Even if the component simple systems are in equilibrium, the compound system as a whole may not be in equilibrium. The thermodynamic space of the compound system is the direct product of the thermodynamic spaces of the constituent simple systems. Just as in the thermodynamic space of a simple system, a point in the thermodynamic space of a compound system may correspond to a nonequilibrium state. We have to specify carefully the interactions among the constituent subsystems.

---

---

145 Whether a given path in the thermodynamic space is reversible or not depends on the context. In the case of thermos, the process is undoubtedly irreversible. However, you could cool your coffee by removing heat reversibly by producing work. The path for your coffee may not be different from the case just above.

145 If we need spatially inhomogeneous states, the system will be partitioned into sufficiently homogeneous macroscopic subsystems; if this is impossible, we will not discuss it thermodynamically in this book.
9.10 Why thermodynamics can be useful
When an initial state and a final equilibrium state are given, the change of a state function between these two states does not depend on the actual process but only on these initial and final equilibrium states. Even if the actual process connecting these two states is not a quasistatic process (i.e., does not lie in the thermodynamic space), we can thermodynamically compute the change of any state function during the process with the aid of an appropriate (appropriately devised) quasistatic process connecting the same end points. This makes thermodynamics extremely useful in practice.

9.11 Thermal contact
Empirically we know that even if there is no exchange of work or matter, two systems in contact can exchange their energies. Such a special contact is called thermal contact. If two systems A and B are in thermal contact and are in equilibrium as a compound system, i.e., when regarded as a single system, we say A and B are in thermal equilibrium.

If the systems A and B are in thermal equilibrium, and if B and C are in thermal equilibrium, then so are the systems A and C. That is, the thermal equilibrium relation is an equivalence relation. This is called the zeroth law of thermodynamics. We can say that this equivalence relation is expressed as the equality relation between the temperatures of the systems.

Traditionally, the existence of (empirical) temperature is apparently deduced from the zeroth law, but actually, the argument is not even water tight. Besides, we do not need the zeroth law.

9.12 The first law of thermodynamics
As we have already noted, the first law of thermodynamics is essentially the conservation of mechanical energy (= internal energy) of the system. Mayer, Joule, Helmholtz and others established that the conservation of mechanical energy implies the existence of a state function $E$ called internal energy. Its change $\Delta E$ cannot be explained solely in terms of the net work $W$ supplied to the system, and the deficit $Q$ is understood as the net heat given to the system:

$$\Delta E = W + Q.$$  \hspace{1cm} (9.1)
Thus, in terms of thermodynamic coordinates that can be understood and quantified solely electrodynamically something called ‘heat’ whose ‘true nature’ is not very clear is macroscopically quantified. Notice that although $E$ is a state function, neither $W$ nor $Q$ is a state function; they depend explicitly on the path connecting the initial and the final equilibrium states (the path may not be in the thermodynamic space).

9.13 Sign convention
Let us make the sign convention explicit. The sign is seen from the system’s point of view: everything imported to the system is positive, and exported negative. For example, if you do work to the system, $W > 0$. If you get some useful work from the system $W < 0$.

9.14 Volume work
When the change is quasistatic, $W$ and $Q$ are determined by the equilibrium states of the system along the quasistatic path.

For example, the work $dW$ required to change the system volume from $V$ to $V + dV$ is given by (see Fig. 9.4)

\[ dW = -Fdl, \]  
(9.2)

if the infinitesimal displacement of the piston is $dl$ and the external force is $F$. Here, the differential expressing the infinitesimal work is written as $dW$ instead of $dW$ to indicate clearly that the infinitesimal is not the differential of a state function (not a perfect differential). According to our sign convention, if the system is done a work, then $dW > 0$, but this happens when the volume shrinks, that is, when $dl < 0$. Therefore, (9.2) has a minus sign. If the volume change is quasistatic, then the system pressure $P$ and the external force $F$ is always in balance,

\[ F = A \times P, \]  
(9.3)
where \( A \) is the cross section of the piston. Thus, we have arrived at the formula applicable to the quasistatic process:

\[
d'W = -PAdl = -PdV. \tag{9.4}
\]

If the process is fast, there would not be sufficient time for the system to equilibrate. For example, when we compress the system rapidly, the force necessary and the force given by (9.3) can be different; even the pressure \( P \) may not be well defined. Consequently, (9.4) does not hold (the work actually done is larger than given by (9.4)).

Thus, although the first law is essentially the conservation of mechanical energy, to write it in terms of a small number of variables, the change must be slow (quasiequilibrium).

### 9.15 Magnetic work

The electromagnetic work can be written as (see Q9-2 at the end of this section)

\[
d'W = H \cdot dM, \tag{9.5}
\]

\[
d'W = E \cdot dP, \tag{9.6}
\]

where \( H \) is the magnetic field, \( M \) the magnetization, \( E \) the electric field, and \( P \) the polarization.

### 9.16 Prehistory of the second law

Joule quantitatively demonstrated (in 1847\(^{146}\)) that work can be converted into heat,\(^{147}\) and believed that work and heat were equivalent, but long before him Carnot (1796-1832) had already established (published in 1824;\(^{148}\)) the impossibility of complete conversion of heat into work. His brother told him (in 1844\(^{149}\)) to pay attention to the work of Carnot. Thomson (later Lord Kelvin) recognized the importance of

---

\(^{146}\)1847: C. Bronte, Jane Eyre, E. Bronte, Wuthering Heights and W. H. Prescott A History of the Conquest of Peru were published.

\(^{147}\)\(\langle\)Work equivalent was established by Mayer\(\rangle\) The work equivalent of heat was obtained by Mayer five years before Joule in 1842 with the aid of Mayer’s cycle in 10.10.

\(^{148}\)1824: Beethoven’s 9th Symphony (Karajan+BPO) premiered; Thomson (also Kirchhoff and Smetana) was born.

\(^{149}\)1844: The first electrical telegram was sent by Morse; Goodyear patented vulcanization; Notice that Turner’s famous Rain, Steam and Speed, The Great Western Railway (National Gallery, London) was this year (see footnote 145 below). Mozart died]
Joule’s work, but since he also accepted Carnot he did not believe work and heat were equivalent (Joule rejected Carnot). Thomson realized that Carnot’s work could establish a universal scale of temperature, and introduced the concept of absolute temperature, but he failed to grasp the real relation between heat and work. To resolve the conflict between Joule and Carnot Thomson looked for further empirical facts, and failed to establish thermodynamics.

Clausius did not think further empirical facts were needed to resolve the conflict between Joule and Carnot. Carnot clearly recognized that only when there are hotter and colder heat baths can we produce work; there is a fundamental asymmetry between heat and work. Clausius understood this as follows: Temperature is the ‘price’ of energy. You cannot simply promote the price of energy (you cannot transfer energy from a colder to a hotter bath). Work corresponds to heat at $T = \infty$. Thus, thermodynamics was established by Clausius.

Now, it seems generally accepted that Clausius and Thomson (independently) constructed thermodynamics. But this is largely due to the British propaganda by Tait. If Great Britain were defeated instead of Germany in WWI, the history would not have been distorted so easily.

Perhaps, a much more important prehistoric fact is that industry was far ahead.

---

150 Peter Guthrie Tait (1831-1901).
151 *Industry was far ahead*. The reader should compare the years in this footnote and those in the main text of this entry. Watt’s steam engines were 1760-70, Trevithick’s steam locomotive (‘Puffing Devil’) was 1804 and Stephenson’s Locomotion for Stockton and Darlington Railway was 1825. Steamboats were earlier. Robert Fulton’s boat with a Watt steam engine was 1807 (between New York and Albany 240 km for 32 hrs) [this year, Beethoven 4th symphony premiered; Jacques-Louis David, *The Coronation of Napoleon*].

Newcomen engine [https://www.youtube.com/watch?v=wZZF07T955s](https://www.youtube.com/watch?v=wZZF07T955s) (this automatically continued to the Watts engine) and

Watts engine [https://www.youtube.com/watch?v=6mNsPjHqzx4](https://www.youtube.com/watch?v=6mNsPjHqzx4)

Good animation of the Newcomen engine: [https://www.youtube.com/watch?v=9GqVQPMcY4](https://www.youtube.com/watch?v=9GqVQPMcY4)

Only working Newcomen’s engine at Black Country Living Museum is [https://www.youtube.com/watch?v=HC6LUWSBjyk](https://www.youtube.com/watch?v=HC6LUWSBjyk)

All above in one: [https://www.youtube.com/watch?v=Q1tRwiiu4UQ](https://www.youtube.com/watch?v=Q1tRwiiu4UQ)

Watt was a versatile inventor: copying machine: [https://www.youtube.com/watch?v=bKERVTlpgMO](https://www.youtube.com/watch?v=bKERVTlpgMO)

Why in Britten in the 18th century? [https://www.youtube.com/watch?v=sWppmWC0Qw](https://www.youtube.com/watch?v=sWppmWC0Qw)

J Hutton and Scottish Renaissance, which is the backdrop of study of heat and Watt’s invention, is described in the following (geology) video (with Kelvin’s blunder) [https://www.youtube.com/watch?v=FYuI2uZLmg](https://www.youtube.com/watch?v=FYuI2uZLmg)

Robert Fulton [https://www.youtube.com/watch?v=2w6xQdswYE](https://www.youtube.com/watch?v=2w6xQdswYE)

Glass Stevenson’s engine [https://www.youtube.com/watch?v=73txXT21aZU](https://www.youtube.com/watch?v=73txXT21aZU)

Replica Rocket [https://www.youtube.com/watch?v=yNnOLC_9imY](https://www.youtube.com/watch?v=yNnOLC_9imY)
9.17 The second law of thermodynamics

The second law of thermodynamics summarizes what Carnot and Clausius clearly understood as follows. Two famous expressions are:

Clausius’ principle: Heat cannot be transferred from a colder to a hotter body spontaneously.

Kelvin’s principle: A process cannot occur whose only effect is the complete conversion of heat into work. (No existence of perpetum mobile of the second kind; there is no engine which can produce work without a radiator.)

Notice that Clausius’ principle contains Kelvin’s principle, if we understand work as the heat energy at $T = \infty$ as Clausius recognized.

Here, we use the second law in the following form:

9.18 Planck’s principle and adiabatic process

![Planck's principle](https://www.youtube.com/watch?v=3woUopc1ZS4)

History up to ‘Rocket.’ [https://www.youtube.com/watch?v=wOGYZC-IJPQ](https://www.youtube.com/watch?v=wOGYZC-IJPQ)

---

Planck’s law: In an adiabatic process if all the work coordinates return to their
original value, $\Delta E \geq 0$.

Here, ‘adiabatic process’ must be explained. In short, it is a process without any exchange of heat with the surroundings (a process realized in a Dewar jar).\textsuperscript{152}

The first law implies adiabatically and quasistatically\textsuperscript{153}

$$dE = \sum_i x_i dX_i,$$

where $(x_i, X_i)$ are conjugate pairs for work coordinates (non-thermal variables). The variables $E$ and $X_i$ span the thermodynamic space (of the system under study).

\textbf{9.19 Planck’s principle, Kelvin’s principle and Clausius’ principle are equivalent}

None is more fundamental than the rest:

If Planck’s principle is violated, then adiabatic work can reduce the system energy. That is, work can be produced by a single heat bath. Therefore, Kelvin’s principle is violated.

If Kelvin’s principle is violated, then we can get work from a cold bath and do work on a hotter bath to increase its energy. Thus, Clausius’ principle is violated.

If Clausius’ principle is violated, then we can convert a uniform system into colder and hotter portions and produce work to make the portions to be the same temperature again by a cyclic change of the work coordinates. Thus, Planck’s principle is violated.

We have demonstrated, symbolically, P $\Rightarrow$ C $\Rightarrow$ K $\Rightarrow$ P. That is, all the principles mentioned here are equivalent.

\textbf{9.20 Entropy was not understood by British scientists}

\textsuperscript{152}There is a special wall called an \textit{adiabatic wall} such that for a system surrounded by this wall the necessary work to bring the system from a given initial equilibrium state to a specified final equilibrium state is independent of the actual process but is dependent only on these two end states of the process. A process that can be realized in a system surrounded by adiabatic walls is an adiabatic process. Furthermore, even if a process is realized without surrounded by adiabatic walls but the same process can be realized surrounded by adiabatic walls, it is an adiabatic process. This turned about to be identical to the process without (net) heat exchange with its environment. Thus, even if a system is attached to a heat bath, a process in the system can be adiabatic. Notice that adiabatic process need not be describable in terms of pure mechanics.

\textsuperscript{153}often ‘quasistatically’ is not mentioned explicitly, but to describe the process in terms differentials of state variables, the process must be sufficiently slow.
Now, we wish to demonstrate that the second law implies the existence of a state function called ‘entropy’ which was introduced by Clausius. In introductory thermodynamics it seems unanimously recognized that entropy is a difficult concept to grasp. This may also be a misconception/misunderstanding spread by the British.

British people resisted to recognize entropy for a while. Even Maxwell, who used entropy correctly for the first time in England, misunderstood it initially. In reality, English speaking scientists were rescued by Gibbs who correctly understood thermodynamics.

In summary, Clausius recognized the following. The constant entropy surfaces foliate the thermodynamic space. These surfaces are called ‘adiabats’ or ‘isentropic hypersurfaces.’ If state A has a larger entropy than state B, then we can never go from A to B adiabatically.

### 9.21 Entropy exists

At least once in your life you should try to reproduce the following explanation of the existence of entropy to your intelligent lay friends.

![Figure 9.6](image)

Figure 9.6: Let Q be a state on a vertical line L (along which we can move with heat exchange alone) that can be reached from state P adiabatically and reversibly. If state A can be reached by an adiabatic process from P, then adiabatically we can go from Q to A via P, violating Planck’s principle. Thus, the shaded portion is inaccessible from P adiabatically. If B can be reached by an adiabatic reversible process from P, then adiabatically we can go from B to Q via P, violating Planck’s principle, again. Thus, Q is unique: there is only one point on L that can be reached from P adiabatically and reversibly. (We can adiabatically go from P to B, but it is an irreversible process.)

Choose an arbitrary point P and a vertical line L in the realizable portion of the thermodynamic space. This line is parallel to the energy axis (all the work coordinates are kept constant), along which we can change the states only by exchanging
heat with the external world. Let us find a quasistatic adiabatic and reversible path\textsuperscript{154} connecting $P$ and $L$.

Suppose the path lands on $L$ at point $Q$. Can we also reach other points on $L$ in the same fashion? Planck tells us $A$ is inaccessible; if possible, we can adiabatically go to $A$ from $Q$, contradicting Planck. If we could go to $B$ adiabatically and reversibly from $P$, then we can go to $Q$ adiabatically via $P$, again contradicting Planck’s principle. Thus, we have learned that the point on $L$ we can reach from $P$ adiabatically and reversibly is only $Q$.

Now, moving the stick $L$ throughout the space keeping it parallel to the energy axis, we can construct a hypersurface consisting of points adiabatically, quasistatically and reversibly accessible from point $P$. This is an \textit{adiabat} containing $P$.

\subsection*{9.22 Entropy stratifies the thermodynamic space}
Adiabats foliate the thermodynamic space. That is, no two different adiabatic surfaces cross each other. See Fig. 9.7 Left to understand that these sheets = adiabats cannot cross; crossing means Planck is violated. This implies that we can define a state function $S$, whose level sets are given by these sheets ($S = \text{constant}$ defines an adiabat).

The adiabats do not have any ‘overhangs.’ As you can see from Fig. 9.7 Right, the reason is the same as that for no crossing.

Thus we have realized that the thermodynamic space is stratified (or foliated) into layers vertically stacked respecting their order, so we can define a state function by appropriately assigning real numbers continuously according to their heights along

\textsuperscript{154}Why is such an awkward description of the path? Reversibility does not logically guarantee quasiequilibrium; quasiequilibrium does not mean reversible. This is the reason. However, intuitively, we may say ‘reversible path,’ because not quasistatic reversible process is not very realistic.
a vertical line; in other words, we can define a state function which is an increasing function of $E$ under constant work coordinates. How can we change the value $S$ of this function?

Obviously, we can change $S$ by going up or down along $L$ while the work coordinates being kept fixed; that is, we can change $S$ by adding or subtracting heat $Q$. Since we have assumed that $S$ increases with $E$, for $d'Q > 0$ we must have $dS > 0$. Therefore, we may define $S$ so that $dS \propto d'Q$ holds. Since $Q$ is extensive,\(^{155}\) so must be $S$.

### 9.23 Entropy and heat

Suppose two systems are in contact through a wall that allows only the exchange of heat (that is, in thermal contact), and they are in thermal equilibrium. Exchange of heat $d'Q$ between the thermally equilibrated systems is a reversible process (say, system I gains $d'Q_I = dQ$ and II $d'Q_{II} = -dQ$), so this process occurs within a single adiabat of the compound system (i.e., the two systems considered together as a single system). If we write $d'Q_X = \theta_X dS_X$ ($X = I$ or $II$), with the aid of the additivity of $S$,

$$0 = dS_I + dS_{II} = d'Q \left( \frac{1}{\theta_I} - \frac{1}{\theta_{II}} \right). \quad (9.8)$$

This implies $\theta_I = \theta_{II}$. That is, when two systems are in thermal equilibrium, the proportionality constants are also the same. Hence, we may interpret the proportionality factor as an empirical temperature (cf. the zeroth law).

The introduced temperature can be chosen as a universal temperature $T$ called the absolute temperature. Hence, in the quasistatic process we can write\(^{156}\)

$$d'Q = T dS. \quad (9.9)$$

---

\(^{155}\) That is, if we double the system, we must double the heat to reach the same thermodynamic state characterized by the same intensive parameters and densities (= extensive variables per volume).

\(^{156}\) Precisely speaking, we must show that this $T$ is identical to the $T$ appearing in the ideal gas law. To this end we have only to consider the Carnot cycle, or to compute the efficiency $\eta$ of an ideal engine. This will be done in the next lecture, but we will obtain $\eta = 1 - \theta_1/\theta_2$ (assuming that $\theta_1 < \theta_2$) [This is Thomson’s fundamental idea to define temperature in a materials-free fashion]. If we use an ideal gas we obtain $\eta = 1 - T_1/T_2$, so $\theta$ and $T$ are identical (up to the choice of units). As we will see ideal gases contradict the third law of thermodynamics, so there are people who assert ideal gases are unphysical and should not be used to develop the basic theoretical framework: Thomson clearly thought particular material should not be used to develop basic theories. However, if pressure is sufficiently reduced, then any real gas becomes an ideal gas however low its temperature is. Therefore, as long as we clearly recognize this condition, there is no fundamental difficulty in using ideal gases to develop fundamental theories.
9.24 Gibbs relation
Now we can write down the infinitesimal version of the first law of thermodynamics for the quasistatic process as follows:

\[ dE = TdS - PdV + \mu dN + B \cdot dM + \cdots. \]  \hspace{1cm} (9.10)

This is called the *Gibbs relation*. The chief concern of thermodynamics up to Gibbs was to formulate the second law and to prove the existence of entropy. Gibbs then utilized entropy and reformulated thermodynamics as a system even practically useful. The very starting point of this new formulation was this relation, which Gibbs wrote down for the first time. Notice that each term consists of a product of a conjugate pair: an intensive quantity and \( d \) (the corresponding (i.e., conjugate) extensive quantity). Also do not forget the minus sign in front of \( PdV \).
Q9-1. An equilibrium state of a macroscopic system can be phenomenologically described by thermodynamics. There is a special set of variables called the thermodynamic coordinates to describe equilibrium states. Briefly explain within 10 lines (with an ordinary letter size, please) what a thermodynamic coordinate system is and why it is special.

Soln.
The thermodynamic coordinate system consists of internal energy $E$ and (extensive) work coordinates describing macroscopic mechanical work that can be done to the system.

1. [Pure mechanical nature] Since $E$ is essentially the mechanical (including electromagnetic) energy of the particles in the system and since work coordinates are described, controlled and measured with the aid of macroscopic mechanics (including electromagnetism), thermodynamic experiments can be described in terms of these coordinates without clear characterization of heat.

2. [Unique specification of equilibrium states] If two states have identical thermodynamic coordinates, then these states cannot be distinguished thermodynamically. Or you can say thermodynamic coordinates specify equilibrium states uniquely.

Q9-2. The work required to increase the magnetization (= the total magnetic moment in the system) of a block in the external magnetic field $H$ is written as $dW = H \cdot dM$. We have not shown this. This is not very trivial, because not all the energy is stored in the block under study; some portion is stored as potential energy in the ‘relation’ between the block and the system creating the external magnetic field. We know that this potential energy is $-H \cdot M$ (probably, you remember that the energy of a magnetic dipole $\mu$ is minimum, when $H$ and $\mu$ are parallel: $E = -\mu \cdot H$).

Since $M = \sum \mu$, where the summation is over all the magnetic dipoles in the block, let us study individual magnetic moment. We assume that the (size of the) magnetic dipole changes due to the magnetic field.

$$F = \mu(H(x)) \frac{dH(x)}{dx}. \quad (9.11)$$
Since we are doing thermodynamics, we must bring the magnetic dipole from infinity slowly to the present position \( x \). To perform such an experiment, you must apply a force opposing the above force (i.e., \(-F\)) while moving the magnetic dipole.

(1) What is the work \( W \) you do to the whole system (the dipole + the bar magnet) while dragging the dipole from \(-\infty\) to \( x \)? [This simply asks your work expenditure. Since the force you exert and the displacement are both given, it is an elementary question.]

(2) However, this energy \( W \) is stored not only in the block containing the magnetic dipoles, but also between the bar magnet and the dipole as the potential energy at \( x \) (as given above). Show that the energy stored in the dipole is

\[
W + \mu(H(x))H(x) = \int_{\mu(0)}^{\mu(H(x))} H(x')d\mu(H(x')).
\] (9.12)

Therefore, \( dE = HdM \) if only the magnetization (the total dipole moment \( \sum \mu \)) is changed among the work coordinates.

**Solt.** (1) The force you exert is \(-F\) (not \( F \); without your application of brake, the ‘block’ would fly to the bar magnet)

\[
W = -\int_{-\infty}^{x} Fdx' = -\int_{0}^{H(x)} \mu(H(x'))dH(x').
\] (9.13)

Here, the dependence of \( \mu \) on \( H \) is explicitly written. This implies that the total work done to the system consisting of the block (containing dipoles) and the bar magnet reads

\[
W = -\int_{0}^{H(x)} M(H)dH.
\] (9.14)

(2) Let us honestly compute this sum (9.12).

\[
-\int_{0}^{H(x)} \mu(H(x'))dH(x') + \mu(H(x))H(x) = -\int_{0}^{H(x)} \mu(H(x'))dH(x') + \int_{0}^{H(x)} d[\mu(H(x'))H(x')]
\]

\[
= -\int_{0}^{H(x)} \mu dH + \int_{0}^{H(x)} d[\mu H] = \int_{0}^{\mu(H(x))} H(x')d\mu(H(x')).
\]

If we sum this over all the dipoles in the block, we get

\[
\int_{M(0)}^{M(H(x))} H(x')dM(H(x')).
\]

Therefore, \( dE = HdM \) if only the magnetization (the total dipole moment \( \sum \mu \)) is changed among the work coordinates.
Q9-3. Suppose we know the following equations of state of a gas:

\[ T = \sqrt{E/V}, \quad P = E/V. \]  

Find \( S = S(E,V) \) up to an additive constant that you cannot fix.

**Soln.**

Let us write down \( dS \) (i.e., the Gibbs relation)

\[ dS = \frac{1}{T} dE + \frac{P}{T} dV. \]

Introducing the given equations of state, we obtain

\[ dS = \frac{V^{1/2}}{E^{1/2}} dE + \frac{E^{1/2}}{V^{1/2}} dV. \]

Since \( S \) is a state function, for example, \( S(E,V) - S(1,1) \) does not depend on the integration paths along which we integrate this differential form from \((E,V) = (1,1)\) to \((E,V)\). Therefore, let us use a convenient path that is piecewise parallel to the coordinate axes. First, let us go from \( E = 1 \) to \( E \) along \( V = 1 \), and then, we go from \( V = 1 \) to \( V \) along \( E = \) constant:

\[ S(E,V) - S(1,1) = \int_1^E \frac{1}{E^{1/2}} dE + \int_1^V \frac{E^{1/2}}{V^{1/2}} dV = 2(E^{1/2} - 1) + 2E^{1/2}(V^{1/2} - 1). \]

In the first integral \( V \) is fixed at 1, and in the second \( E \) is fixed at its final value.

Thus, we have obtained

\[ S(E,V) = 2E^{1/2}V^{1/2} + \text{const}. \]

Notice that ‘miraculously’ the term dependent on the starting position is cleanly separated out as a constant term. [This is, of course, guaranteed by a Maxwell’s relation. You’d better check this.]

Q9-4. On the \( PV \)-diagram of any gas, two adiabatic curves (such as corresponding to \( PV^\gamma = \) const. for the ideal gas) never cross. Demonstrate this (see Fig. 9.9; the situation depicted in this figure never happens). [This is almost a trivial question.]

**Soln.**

Look at the ‘triangular’ region surrounded by a high temperature isotherm and two adiabats. If we go around this cycle in the clockwise direction

\[ - \oint PdV = -(\text{the area surrounded by these three curves}) < 0, \]

which is the work gained by the gas, but it is negative. That is, it can perform work using a single heat bath, violating Kelvin’s law.
Figure 9.9: Black curves are isotherms (\( T \) const), and red curves are adiabats (\( S \) const). In this diagram, two adiabats cross, but such crossing never occurs.
10 Thermodynamics: General consequences

Summary
* What you should remember about thermodynamics is summarized.
* Under an adiabatic condition, spontaneous changes imply increasing entropy.
* Under an adiabatic condition, a system reaches equilibrium when its entropy becomes maximum (the principle of maximum entropy).
* There is a general logic to extend the results for adiabatic systems to non-adiabatic systems. This gives you Clausius’ inequality.
* The crux of thermodynamic computation is to devise quasistatic processes.
* The efficiency to convert heat into work is bounded by a maximum value determined by the temperatures of the heat sources (Carnot’s theorem).

Key words
Clausius’ inequality, entropy maximization principle, equilibrium conditions, reversible engine, Carnot’s theorem, heat pump

What you should be able to do
* To compute entropy changes for simple processes.
* Remember the key features of the ideal gas.
* To be able to compute the efficiencies of an ideal engine.
* To show that the reversible engine is the best engine.

10.1 Summary of basic principles

The basic laws of thermodynamics are the summary of the experiences of us macroscopic organisms ([N] indicates the closest traditional ‘N-th law’):

[O] There is a state called an equilibrium state. Equilibrium states of a system may be described in terms of thermodynamic coordinates \((E, X_i)\), where \(E\) is the internal energy and \(X_i\) are work coordinates. The equilibrium state exhibits partitioning-

\[157 \text{[Nernst’s joke on the three principles]}\] Kurt Mendelssohn writes, “When lecturing on ‘his’ heat theorem, Nernst was careful to point to an interesting numerical phenomenon concerning the discovery of the three fundamental laws of thermodynamics. The first one had three authors, Mayer, Joule and Helmholtz; the second had two, Carnot and Clausius; whereas the third was the work of one man only, Nernst. This showed conclusively that thermodynamics was now complete since the authorship of a hypothetical fourth law would have to be zero.” (The world of Walther Nernst: the rise and fall of German Science 1864-1941 (ebook form from Plunket Lake Press, 2015; the original 1973) Chapter 4.
rejoining invariance (9.3).

[I] The conservation of energy: \( \Delta E = Q + W \), or for infinitesimal changes \( dE = dQ - PdV + BdM + xdx \).
\( \text{158} \) the variables appear in ‘conjugate pairs’: \((-P, V), (B, M), (x, X)\) (for a generic pair), etc. (9.12).

[II] The thermodynamic space is foliated into \( S = \text{constant} \) (hyper)surfaces. With adiabatic quasistatic (thus reversible) processes we cannot get out of a given \( S = \text{const.} \) surface. With work only, \( \Delta S < 0 \) never happens; to reduce entropy we definitely need cooling. For a quasistatic process the Gibbs relation holds: \( dE = TS \ dS + \sum_i x_i dX_i \). Often \( dS = \frac{1}{T} dE + \frac{\mu}{T} dV - \frac{x}{T} dX + \cdots \) is convenient. (20.3)

[III] \( S = 0 \) in the limit \( T \to 0 \). This is the third law we will encounter in 16.4.

[IV] Thermodynamic variables are either extensive or intensive. The total amount of an extensive quantity of a compound system is the sum of the extensive quantities of the subsystems (additivity) (9.4).

Practically,
(i) Thermodynamics can be used to compute the state function change caused by any process connecting an initial equilibrium state A and a final equilibrium state B.
(ii) To this end we devise a convenient quasistatic path from A to B in the thermodynamic space along which we can use the Gibbs relation mentioned in (II) above.

10.2 Entropy maximization principle
Entropy cannot be reduced by any adiabatic process. Therefore, if an equilibrium state changes into another equilibrium state through modification of only the work coordinates under an adiabatic condition,\( \text{159} \) the entropy of the system generally increases.

Suppose the initial system is in equilibrium but with some constraints (say, compartmentalized with internal walls). If we remove the constraints, the system would evolve to a new equilibrium state (Fig. 10.1).

Since the change is spontaneous, generally, this final state has a larger entropy. This

\( \text{158} \) (Standard state function symbols) We stick to the standard notations:
\( E \): internal energy, \( S \): entropy, \( T \): temperature, \( P \): pressure, \( V \): volume, \( B \) (\( B \) or \( h \)): magnetic field, \( M \) (\( M \)): magnetization, \( \mu \): chemical potential, \( N \): number of particles. We use \( X \) for a generic work coordinate (extensive quantity) and \( x \) for its intensive conjugate (with respect to energy).

\( \text{159} \) ‘Adiabatic’ implies no exchange of heat. Then, the reader may infer that thermal contact with a single heat bath is admissible if there is no net heat exchange. This is correct. Notice that ‘adiabatic condition’ does not mean \( R \to 0 \) (external noise zero) limit (Section 12).
Figure 10.1: Initially, assume that the system is in equilibrium in the presence of a wall, which may be understood as a symbol of a certain constraint. The total entropy of this system as a whole (i.e., as a compound system 9.9) is $S_1 + S_2$. The whole system is under an adiabatic condition. When the wall is removed (i.e., the constraint is removed), the system evolves to a new equilibrium state with a larger entropy spontaneously (irreversibly). That is, the final entropy $S$ must satisfy $S \geq S_1 + S_2$ (the principle of increasing entropy).

is the principle of increasing entropy.\footnote{However, this does not claim the system entropy increases at every intermediate time during the evolution process, because thermodynamic entropy is defined only for equilibrium states.}

A spontaneous change in an adiabatic system increases its entropy, so if the system entropy reaches the maximum under a given constraint, the system reaches its equilibrium state under the constraint. This is the entropy maximization principle.\footnote{\textbf{Remark on entropy max principle}} Thus, the change $\delta S$ of the system entropy due to any virtual change (perturbation) of the system tells us that (stability and evolution criteria):

\begin{align}
\delta S < 0 & \iff \text{the state is thermodynamically stable}, \quad (10.1) \\
\delta S > 0 & \iff \text{the state spontaneously evolves}. \quad (10.2)
\end{align}

Therefore, the second law gives us a variational principle in terms of entropy to find a stable equilibrium state for an adiabatic system.

In the above description of the stability criterion, we mentioned 'virtual changes or perturbations'.\footnote{\textbf{Remark on entropy max principle}} In reality, however, they are not virtual in most cases, but are...
actually produced spontaneously by thermal fluctuations. Thus, as long as thermal fluctuations are not suppressed, whenever the system entropy can increase, the system evolves to maximize its entropy; behind any variational principle are fluctuations to substantiate it.

### 10.3 Entropy is concave

Let us join two systems made of the same substances to make a single system. The entropy maximization principle tells us that the entropy of the resultant compound system is given by

\[
S(E, X) = \max[S(E_1, X_1) + S(E_2, X_2)],
\]

where the maximum is taken over all the partitions of \( E \) and \( X \) between the two systems as \( E = E_1 + E_2 \) and \( X = X_1 + X_2 \) (\( X \) collectively denotes work coordinates). This implies with the aid of the extensivity of \( S \) (i.e., \( S(\alpha E, \alpha X) = \alpha S(E, X) \))

\[
S((1 - \alpha)E + \alpha E', (1 - \alpha)X + \alpha X') \geq (1 - \alpha)S(E, X) + \alpha S(E', X')
\]

for any \( \alpha \in [0, 1] \). That is, \( S \) is a concave function (its graph is convex upward) of all the thermodynamic coordinates (see Fig. 10.2A). This implies that the local stability criterion (10.1) holds globally as well.

![Figure 10.2](image)

**Figure 10.2:** An example of the concave function A and that of the convex function B are illustrated; the resultant inequality for B is called Jensen’s inequality. The black dots correspond to the right-hand sides of (10.4) and (10.8), respectively.

### 10.4 Internal energy minimization principle

The entropy maximization principle implies for any deviation \( \Delta X \) of \( X \) from the
equilibrium value\textsuperscript{163}

\[ S(E, X_{eq}) - S(E, X_{eq} + \Delta X) \geq 0. \]  

(10.5)

Therefore, since \( S \) is an increasing function of energy, we can increase the internal energy \( E \) in the second term to \( E' \geq E \) under the \( X = X_{eq} + \Delta X \) condition to satisfy

\[ S(E, X_{eq}) - S(E', X_{eq} + \Delta X) = 0. \]  

(10.6)

This implies that under the constant entropy condition, if an extra constraint to fix \( X \) at \( X_{eq} + \Delta X \) is removed, then the internal energy surely decreases in equilibrium, since \( E \leq E' \). That is, if the internal energy is minimized under a constant entropy condition, the system must be in equilibrium.

**10.5 Internal energy is convex**

Let us join two systems made of the same substances to make a single system. The energy minimization principle tells us that the internal energy of the resultant compound system is given by

\[ E(S_1 + S_2, X) = \min[E(S_1, X_1) + E(S_2, X_2)], \]  

(10.7)

where the minimum is taken over all the partitions of \( S \) and \( X \) between the two systems as \( S = S_1 + S_2 \) and \( X = X_1 + X_2 \). This implies with the aid of the extensivity of \( E \)

\[ E((1 - \alpha)S + \alpha S', (1 - \alpha)X + \alpha X') \leq (1 - \alpha)E(S, X) + \alpha E(S', X') \]  

(10.8)

for any \( \alpha \in [0, 1] \). That is, \( E \) is a convex function (its graph is convex downward) of all the variables (= entropy and work coordinates) (see Fig. 10.2B).

Let us extend our inequalities for thermally isolated systems to thermally non-isolated systems. The following argument exemplifies a standard strategy that we use repeatedly throughout statistical thermodynamics.

**10.6 Extension to non-adiabatic systems**

Let us extend our inequalities for thermally isolated systems to thermally non-isolated systems. The following argument is a standard strategy that we use repeatedly throughout statistical thermodynamics. To consider a system which is not

\textsuperscript{163}Since entropy is defined only for equilibrium states, this means, precisely speaking, that if, with some constraints, we make a new equilibrium state with \( X + \Delta X \) and \( E \), then (10.5) holds.
isolated, that is, a system which is interacting with its environment, we construct a bigger isolated system composed of the system itself (I) and its interacting environment (II) (Fig. 10.3). We assume that both systems are macroscopic, so we may safely ignore the surface effect.

Figure 10.3: The system II is the environment for the system I we are interested in. II is sufficiently large so even if changes in I are irreversible, II remains infinitesimally close to equilibrium (i.e., any change in I causes a quasistatic change in II).

The environment is stationary (in equilibrium), whose intensive thermodynamic variables such as temperature are kept constant. To realize this we take a sufficiently big system (called a reservoir like a thermostat or a chemostat) as the environmental system II. Even if a change is a rather drastic one for the system I itself, it would be negligible for the system II, because it is huge. Therefore, we may assume that any process in the system I is a quasistatic process for system II.

10.7 Clausius’ inequality
The entropy change of the compound system I+II is given by the sum of the entropy change of the system I denoted by $\Delta S_I$ and that of the environment II denoted by $\Delta S_{II}$. Since the whole system I + II is adiabatic, a natural process occurring in the whole system must satisfy

$$\Delta S_I + \Delta S_{II} \geq 0. \quad (10.9)$$

Let $Q (> 0)$ be the heat transferred to the system I from the environment II. From our assumption, we have

$$\Delta S_{II} = -Q/T_e, \quad (10.10)$$

where $T_e$ is the temperature of the environment (system II). The minus sign is because

164 Usually, the amount of heat transferred from a system is obtained from its temperature and heat capacity. Therefore, you might claim that if a system is huge, it would be impossible to measure its temperature change accurately, so consequently heat $Q$ transferred may not be accurately determined. In reality, we can use a thermometer and electric heater to construct a surface (thermostat) that is maintained very accurately at a given temperature and we can measure the needed electricity to maintain it to obtain $Q$. Thus, virtually we can realize an ideal heat bath.
II is losing heat to I. Combining (10.9) and (10.10) yields the following inequality:

$$\Delta S_I \geq \frac{Q}{T_e}. \quad (10.11)$$

This is *Clausius’ inequality* for non-adiabatic systems. This tells us when something spontaneously happens with heat exchange allowed, the actual entropy change is larger than that due to a reversible process. Of course, for adiabatic systems $Q$ vanishes, so we recover the principle of maximum entropy 10.2.

If the change in I is reversible, then $\Delta S_I = Q/T_e$ should hold; (10.11) implies that ‘excessive entropy’ has been produced in I by the irreversibility of the process.

### 10.8 Equilibrium conditions between two systems

As an application of the entropy maximization principle, let us study the equilibrium conditions for two systems I and II interacting through various walls.

![Figure 10.4: The thick vertical segment denotes the wall that selectively allows the exchange of a certain extensive quantity.](image)

i) Consider a rigid impermeable wall which is diathermal. Thus, the two systems in contact through this wall exchange energy (internal energy) in the form of heat. The total entropy of the system $S$ is the sum of the entropy of each system $S_I$ and $S_{II}$. The total internal energy $E$ is also the sum of subsystem internal energies $E_I$ and $E_{II}$ (extensivity). We isolate the compound system and ask the equilibrium condition for the system. We should maximize the total entropy with respect to the variation of $E_I$ and $E_{II}$:

$$\delta S = \frac{\partial S_I}{\partial E_I} \delta E_I + \frac{\partial S_{II}}{\partial E_{II}} \delta E_{II} = \left( \frac{\partial S_I}{\partial E_I} - \frac{\partial S_{II}}{\partial E_{II}} \right) \delta E_I = 0, \quad (10.12)$$

where we have used that $\delta E = 0$ or $\delta E_I = -\delta E_{II}$. Hence, the equilibrium condition is

$$\frac{\partial S_I}{\partial E_I} = \frac{\partial S_{II}}{\partial E_{II}}, \quad (10.13)$$

or $T_I = T_{II}$.

ii) Consider a diathermal impermeable wall which is movable. In this case the two
systems can exchange energy and volume. If we assume that the total volume of the system is kept constant, the equilibrium condition should be
\[
\delta S = \frac{\partial S_I}{\partial V_I} \delta V_I + \frac{\partial S_{II}}{\partial V_{II}} \delta V_{II} = \left( \frac{\partial S_I}{\partial V_I} - \frac{\partial S_{II}}{\partial V_{II}} \right) \delta V_I = 0, \tag{10.14}
\]
and \( T_I = T_{II} \), that is,
\[
\frac{\partial S_I}{\partial V_I} = \frac{\partial S_{II}}{\partial V_{II}} \tag{10.15}
\]
and \( T_I = T_{II} \). Therefore, \( P_I = P_{II} \) is also required.

**Remark** If the wall is adiabatic, then it cannot exchange heat, so there is no way to exchange entropy. This suggests that to use the Gibbs relation (9.10) directly is convenient. \( P_I = P_{II} \) is the condition; we cannot say anything about the temperatures.

### 10.9 Equilibrium system has heat bath contact with which does not alter state

It is almost never emphasized but perhaps the most important characteristic of an equilibrium macrosystem is that there is always a heat bath contact with which does not alter the equilibrium system. I call it the private (or individual) heat bath for the state.

Notice that a system attached to a heat bath cannot be described by any mechanics.\(^{165}\) Thus, we may say, in equilibrium at least, thermodynamics transcends mechanics (quantum or not). Following the 19th and the 20th century tradition, we still believe smaller scales are more basic without firm empirical supporting arguments. Of course, this point of view may well be the correct way even empirically to understand our world, we should be skeptical to be faithful to the fundamental of science.

#### Thermodynamics through examples

Let us get familiar with thermodynamics through basic practice problems:

**10.10 Mayer’s relation**
Demonstrate *Mayer’s relation*: \( C_P = C_V + R \), where \( C_P \) is the constant pressure

\(^{165}\)The traditional approach is to describe the heat bath as a much bigger isolated purely mechanical system. However, we must not forget that the larger the system the harder to maintain it in isolation. Of course, you cannot enclose the system with a bigger pure mechanical system.
molar specific heat and \( C_V \) the constant volume molar specific heat of an ideal gas.

First of all, we must identify the quantities in terms of thermodynamic variables. The specific heat under constant \( V \) and constant \( P \) are defined as

\[
C_V = \left. \frac{\partial Q}{\partial T} \right|_V, \quad C_P = \left. \frac{\partial Q}{\partial T} \right|_P.
\]

The first law tells us \( dE = d'Q - PdV \), so

\[
C_V = \left. \frac{\partial E}{\partial T} \right|_V, \quad C_P = \left. \frac{\partial E}{\partial T} \right|_P + P \left. \frac{\partial V}{\partial T} \right|_P.
\]

For an ideal gas \( E \) is dependent only on \( T \) (recall that \( E \) is the kinetic energy of unhindered molecular motion for idea gases), so \( dE = C_V dT \). \( V = RT/P \), so

\[
C_P = \left. \frac{\partial E}{\partial T} \right|_P + P \left. \frac{\partial V}{\partial T} \right|_P = C_V + R.
\]

Mayer obtained this relation with the aid of Mayer’s cycle (Fig. 10.5). Recall that ideal gas has only kinetic energy which is uniquely determined by temperature as \( E = C_V T \).

![Figure 10.5: Mayer’s cycle consists of isobaric compression 1, constant volume heating 2, and adiabatic free expansion (recall the law of constant temperature due to Gay-Lussac 3.](image)

Notice that 3 in Fig. 10.5 does not change \( E \), so for the ideal gas A and C are at the same temperature, say, \( T_2 \). Let the temperature at B be \( T_1 \). The work supplied by the isobaric compression process 1 is \( W = P_1(V_2 - V_1) \). The heat is discarded during this process simultaneously: \( C_P(T_2 - T_1) \). During the process 2 heat \( C_V(T_2 - T_1) \) is absorbed. Therefore, for the cycle as a whole, we have

\[
0 = P_1(V_2 - V_1) + C_P(T_1 - T_2) + C_V(T_2 - T_1) = R(T_2 - T_1) + C_P(T_1 - T_2) + C_V(T_2 - T_1).
\]

That is, \( R - C_P + C_V = 0 \).
10.11 Poisson’s relation
Show that along an adiabatic quasistatic path $PV^\gamma = \text{const.}$, where $\gamma = C_P/C_V$. This is called Poisson’s relation.

The first law implies $dE = -PdV$ (adiabatic and quasistatic!). Also $dE = C_VdT$ (ideal gas). Therefore,

$$0 = C_VdT + PdV = C_Vd(PV/R) + PdV = (C_V/R + 1)PdV + (C_V/R)VdP. \quad (10.20)$$

That is, $\gamma d\log V + d\log P = 0$ with the aid of Mayer’s relation.

10.12 Reversible engine: Carnot’s theorem
Obtain the efficiency $\eta$ (see (10.22) for the definition) of a reversible heat engine, and demonstrate that there is no engine more efficient than the reversible engine (Carnot’s theorem).

A heat engine is a device that absorbs heat from a high temperature heat bath (temperature $T_H$) and converts a portion into work. The remaining energy is discarded to a low temperature heat bath (temperature $T_L$). See Fig. 10.6. Let $Q_H$ and $Q_L$ be the heats the engine absorbs from the high and low temperature heat baths, respectively, per one cycle, and $W$ the work the engine obtains per one cycle (we expect $Q_H > 0$, $Q_L < 0$ and $W < 0$). The first law implies (since the engine does not produce energy)

$$W + Q_H + Q_L = 0. \quad (10.21)$$

The efficiency of an engine is the ratio of the work the engine produces (the benefit we get) to the heat it absorbs from the high-temperature reservoir (the expenditure we pay). Therefore, we define the engine efficiency as (be careful with the sign convention)

$$\eta \equiv \frac{|W|}{Q_H} = -\frac{W}{Q_H} = 1 + \frac{Q_L}{Q_H}. \quad (10.22)$$

150
Let $\Delta S_H$ be the entropy increase of the engine in a single cycle due to the import of heat from the high-temperature bath, and $\Delta S_L$ due to the import of heat from the low-temperature bath. Clausius’ inequality (10.11) tells us that

$$\Delta S_H \geq \frac{Q_H}{T_H}, \quad \Delta S_L \geq \frac{Q_L}{T_L}.$$  \hspace{1cm} (10.23)

Since the engine returns to the original state after a single cycle, $\Delta S = \Delta S_H + \Delta S_L = 0$:

$$0 = \Delta S_H + \Delta S_L \geq \frac{Q_H}{T_H} + \frac{Q_L}{T_L} \Rightarrow \frac{Q_H}{T_H} \leq -\frac{Q_L}{T_L},$$  \hspace{1cm} (10.24)

which implies $T_L/T_H \leq -Q_L/Q_H$. Thus, we have

$$\eta = 1 + \frac{Q_L}{Q_H} \leq 1 - \frac{T_L}{T_H}.$$  \hspace{1cm} (10.25)

If the engine is reversible, it attains the maximum efficiency $\eta = 1 - T_L/T_H$. This statement is called Carnot’s theorem.\textsuperscript{166}

Thomson saw here a possibility of introducing the universal temperature scale based solely on the thermodynamic principles free from any materials; he reached the concept of the absolute temperature in terms of the maximum efficiency.

10.13 The original Carnot’s argument using the Carnot cycle of an ideal gas.\textsuperscript{167}

Carnot conceived the following engine (the Carnot engine) which utilizes an ideal gas (in this exposition, 1 mole of it) as its working substance (Fig. 10.7). This original demonstration of Carnot’s theorem is much harder than the one we just saw, but may be a good elementary thermodynamics exercise:

(i) The engine does work through expansion while absorbing heat from the high temperature heat source (at $T_H$) (A$\rightarrow$B in Fig. 10.7).

(ii) Then, it continues to expand while doing work and cools from $T_H$ to $T_L$ (B$\rightarrow$C). Notice that this portion was Watt’s novelty.

(iii) Next, the engine volume isothermally shrinks (i.e., some positive work is done to the engine) while discarding heat to the low temperature heat source at $T_L$ (C$\rightarrow$D).

(iv) Finally, the system is compressed adiabatically (again some positive work is done to the engine) and its temperature goes up from $T_L$ to the original $T_H$ (D$\rightarrow$A).

The work added to the system (engine) is

$$W = -\oint_{ABCD} PdV,$$  \hspace{1cm} (10.26)

\textsuperscript{166}If $T_H = +\infty$, then the reversible efficiency is 1. Recall 9.17, according to Clausius, that work is heat from a bath at $T = \infty$.

\textsuperscript{167}The actual original argument due to Carnot relied on the caloric theory (which regarded heat as a substance called caloric), so the exposition given here cannot literally be his original argument, but a correct transliteration was done by Clausius. We need this demonstration to identify the absolute temperature introduced by the ideal gas law and $\theta$ we introduced to relate heat and entropy change in Section 9.
Figure 10.7: The Carnot cycle: AB and CD are quasistatic isothermal processes, and BC and DA are quasistatic adiabatic processes. BC is the key element of Watt’s engine. The working substance is an ideal gas, so during the isothermal process its internal energy is constant. This implies that during isothermal processes the work the system does (or is done to the system) and the heat it absorbs (or it discards) must be identical. Understanding the Carnot engine with the $PV$-diagram was originally due to Clapeyron (1834, thus this diagram is called Clapeyron’s graph), who advocated Carnot’s work.

so it is equal to $(-1) \times$ the area surrounded by the warped quadrangle ABCD in Fig. 10.7. That is, the work done by the engine during its one cycle is the area of ABCD.

During the isothermal process $A \rightarrow B$ the engine does some work to the environment, but its internal energy is constant, because this is an isothermal process for an ideal gas; the work must be paid by the heat $Q_H$ absorbed from the high temperature heat source at $T_H$. Therefore, (notice $dE = d'Q - PdV = 0$)

$$Q_H = \int_{A \rightarrow B} PdV = \int_{A \rightarrow B} \frac{RT_H}{V}dV = RT_H \log \frac{V_B}{V_A} > 0.$$  \hspace{1cm} (10.27)

Similarly, during the isothermal process $C \rightarrow D$ the heat $|Q_L|$ discarded (i.e., $Q_L < 0$) absorbed by the system to the low temperature heat source at $T_L$ must be identical to the work done to the system, so we have

$$|Q_L| = -\int_{C \rightarrow D} PdV = -\int_{C \rightarrow D} \frac{RT_L}{V}dV = RT_L \log \frac{V_C}{V_D}. \hspace{1cm} (10.28)$$

To relate these two formulas, we need the ratios of the volumes related by quasistatic adiabatic processes. Poisson’s relation 10.11 $PV^\gamma = \text{const.}$ implies $TV^\gamma = \text{const.}$ Consequently, $T_HV_A^{\gamma-1} = T_LV_D^{\gamma-1}$ and $T_HV_B^{\gamma-1} = T_LV_C^{\gamma-1}$ hold. This implies that $T_H/T_L = V_D^{\gamma-1}/V_A^{\gamma-1} = V_C^{\gamma-1}/V_B^{\gamma-1}$, or $V_B/V_A = V_C/V_D$. Using this relation in (10.27) and (10.28), we obtain the equality in (10.24). The rest is identical to the argument above, and we get $\eta = 1 - T_L/T_H$.

Carnot’s original proof of his theorem went as follows. Suppose we have an engine B better (more efficient) than the reversible engine R, which can be driven backward by supplying work. Let us drive the reversible engine R backward with engine B and use R as a ‘heat pump’ (see Fig. 10.8).

$|W'| > |W|$, so if we use the output of the ‘better engine’ to drive the reversible engine, we can still utilize the work $|W'| - |W|$. Since the net heat imported to the two engines from
Figure 10.8: The reversible engine R (Left) is now used as a heat pump, and is driven by a (imaginary) better engine B (Right) that can produce work $|W'| > |W|$ using the identical heat sources.

the hotter bath is zero, inevitably, $|Q'_L| < |Q_L|$. That is, $|Q_L| - |Q'_L|$ is absorbed from the colder bath. This implies that work has been extracted from a single heat bath, violating Kelvin’s principle. Hence, there cannot be any better engine than the reversible engine.

### 10.14 Ideal gas: thermodynamic equation of state

The thermodynamic space of an (1 mole) ideal gas is spanned by internal energy $E$ and volume $V$. Compute the entropy difference between the initial equilibrium state $(E_1, V_1)$ and the final equilibrium state $(E_2, V_2)$ for a 1 mole of ideal gas.

Since entropy is a state function, to compute its change between two equilibrium states, we may invent a convenient process connecting these two states. The process we can compute in detail is a quasistatic process.

The first law (or the Gibbs relation) tells us along a quasistatic process

$$dS = \frac{1}{T}dE + \frac{P}{T}dV.$$  \hspace{1cm} (10.29)

Since for a (1 mole) ideal gas $E = C_V T$ and $PV = RT$,

$$dS = \frac{C_V}{E}dE + \frac{R}{V}dV = C_V d\log E + Rd\log V.$$  \hspace{1cm} (10.30)

$dS$ is a perfect differential, so we have only to integrate this along a convenient path (this is the meaning of inventing a convenient process):

$$S(E_2, V_2) = S(E_1, V_1) + C_V \log \frac{E_2}{E_1} + R \log \frac{V_2}{V_1}.$$  \hspace{1cm} (10.31)

In contrast to the usual equation of state $PV = RT$, the relation (which should be called the true equation of state) $S = S(E, V)$ gives you ‘everything’ you wish to know about the ideal gas:

$$P \frac{1}{T} = \frac{\partial S}{\partial V} \bigg|_E = R \frac{1}{V}, \quad \frac{1}{T} = \frac{\partial S}{\partial E} \bigg|_V = \frac{C_V}{E}.$$  \hspace{1cm} (10.32)
Poisson’s relation $PV^\gamma = \text{const.}$ must imply $\Delta S = 0$. Since we derived Poisson’s relation assuming $\Delta S = 0$, this should be, but in any case, let us check this. $P = RT/V = RE/C_VV$, so Poisson’s relation implies $EV^{\gamma-1} = \text{const.}$ for an adiabatic quasistatic process. If this relation holds, then, since $R = C_P - C_V$, the system entropy does not change:

$$S(E_2, V_2) = S(E_1, V_1) + C_V \left[ \log \frac{E_2}{E_1} + (\gamma - 1) \log \frac{V_2}{V_1} \right] = S(E_1, V_1).$$  \hspace{1cm} (10.33)

or we may write with an appropriate base point $E_0$ and $V_0$ in the thermodynamic space

$$S(E, V) = S(E_0, V_0) + C_V \log(E/E_0) + R \log(V/V_0).$$  \hspace{1cm} (10.34)

10.15 Two blocks at different temperatures
There are two blocks with the same heat capacity $C$ at temperature $T_L$ and at $T_H$, respectively. If we bring these blocks to thermal equilibrium reversibly or irreversibly, what is the final common temperature $T_F$ (Fig. 10.9)?

![Diagram of two blocks at different temperatures](image)

Figure 10.9: Initially, two blocks have different temperatures. What is the common temperature $T_F$ when the blocks reach a thermal equilibrium?

If we make a thermal contact between them (assume that the system as a whole is thermally isolated = under an adiabatic condition), a ‘perfectly’ irreversible process occurs, and the final temperature is $T_F = (T_L + T_H)/2$ due to the first law and the result we obtained above. Needless to say, the final entropy of this system must be larger than the initial one, i.e., $\Delta S > 0$ ($\Delta$ always means ‘final’ – ‘initial’). To use thermodynamics, we must invent a quasistatic process connecting the initial and the final equilibrium states. We bring one block from $T_L$ to $T_F$, and the other from $T_H$ to $T_F$ quasistatically, and then join these two. This last step does not change anything. Let us study each block separately.

An important observation is that if the heat exchange is across infinitesimal temperature difference $dT$, then the heat transfer is a quasistatic process (no increase
Therefore, we may prepare numerous heat baths at various temperatures, and use them appropriately in turn to change the temperature of the block gradually (quasistatically). Along this process, we may use thermodynamics. Since $dQ = CdT$, $dS = CdT/T$:

$$\Delta S_1 = \int_{T_L}^{T_F} \frac{CDT}{T} = C \log \frac{T_F}{T_L}. \quad (10.35)$$

We can perform quite an analogous calculation for the other box, so combining the answers, we get

$$\Delta S = \Delta S_1 + \Delta S_2 = C \log \left( \frac{T_F^2}{T_L T_H} \right) = 2C \log \frac{T_F}{\sqrt{T_L T_H}}. \quad (10.36)$$

Here, $T_F = (T_L + T_H)/2$. As can be seen from Fig. 10.10,

$$\Delta S = 2C \left[ \log \left( \frac{T_L + T_H}{2} \right) - \frac{\log T_L + \log T_H}{2} \right] > 0. \quad (10.37)$$

![Figure 10.10: Star denotes $\log[(T_L + T_H)/2]$ and square denotes $\log T_L + \log T_H$, demonstrating $(T_L + T_H)/2 > \sqrt{T_L T_H}$.](image)

(Reversible case) $(10.36)$ implies that if $T_F = \sqrt{T_L T_H}$, then $\Delta S = 0$. There must be a reversible process to realize this. How can you do this? Notice that the internal energy of the system is not conserved:

$$\Delta E = 2CT_F - (CT_L + CT_H) = 2C \left( \sqrt{T_L T_H} - \frac{T_L + T_H}{2} \right) < 0. \quad (10.38)$$

Indeed this $|\Delta E|$ must be exported; the system can (must) do work.

To realize this reversible process we can set up a reversible engine between the
two blocks and operate it until there is no temperature difference. Let us assume that $T'_H$ is the temperature of the hotter block, and $T'_L$ that of the colder block at some time point. Since by operation of the engine, the block temperatures change, so we analyze the engine working when the hotter block temperature is between $T'_H$ and $T'_H + dT'_H$ (notice that $dT'_H < 0$). The entropy change must be zero (a reversible engine) during this temperature change:

$$dS = \frac{dQ_H}{T'_H} + \frac{dQ_L}{T'_L} = C\frac{dT'_H}{T'_H} + C\frac{dT'_L}{T'_L} = 0. \quad (10.39)$$

This implies $d\log(T'_HT'_L) = 0$ or $T'_L/T'_H = $ constant. That is, $T'_F = T'_HT'_L = T_LT_H$, or, as we know, the final temperature must be $T_F = \sqrt{T_LT_H}$.

We know the efficiency of the reversible engine, so

$$-\frac{d'W}{d'Q_H} = 1 - \frac{T'_L}{T'_H}, \quad (10.40)$$

or

$$-d'W = \left(1 - \frac{T'_L}{T'_H}\right) d'Q_H. \quad (10.41)$$

Here, $d'Q_H = -CdT'_H$, because decrease of $T'_H$ implies positive $d'Q_H$. Therefore,

$$-d'W = -\left(1 - \frac{T'_L}{T'_H}\right) CdT'_H = \left(1 - \frac{T'_L}{T'_H}\right) T'_H \left(1 - \frac{T'_L}{T'_H}\right) C(-dT'_H). \quad (10.42)$$

That is, the work we can take out from the system is (note that $T_F = \sqrt{T_LT_H}$)

$$-\Delta W = C(T_L - T_F) + CT_H - C\frac{T_LT_H}{T_F} = 2C \left(\frac{T_L + T_H}{2} - \sqrt{T_LT_H}\right). \quad (10.43)$$

This is positive as shown before. Of course, this is a stupid way to compute $\Delta W$; the answer is obvious from the first law. Trust thermodynamics.

In any case a (great) lesson is: if $\Delta S = 0$, there must be a reversible process to realize the change. Devise it.

We will see such an example in the next lecture.
Q10-1. One mole of an ideal gas expands adiabatically against the external constant pressure $P_0$. The gas is initially at temperature $T_1$ and the pressure is $2P_0$. After expansion the gas reaches a final equilibrium state with the pressure $P_0$. You may assume that the constant volume specific heat $C_V$ is constant throughout the process.

1) What is the final temperature?

2) What is the entropy change $\Delta S$ of this process?

**Soln.**

1) The work done by the gas is $P_0(V_2 - V_1)$, where $V_1$ is the initial volume and $V_2$ is the final volume. The final pressure must be $P_0$, so $2P_0V_1 = RT_1$ and $P_0V_2 = RT_2$. Since the system is adiabatic,

$$\Delta E = -P_0(V_2 - V_1) = RT_1/2 - RT_2.$$ 

The internal energy of an ideal gas is directly related to its temperature, so

$$\Delta E = C_V(T_2 - T_1).$$

These formulas imply

$$RT_1/2 - RT_2 = C_V(T_2 - T_1).$$

Therefore,

$$T_2 = \frac{R + 2C_V}{2(R + C_V)} T_1.$$ 

2) To compute $\Delta S$ we need $T_2/T_1$ and $V_2/V_1$.

$$\frac{V_2}{V_1} = \frac{T_1/2P_0}{T_2/P_0} = \frac{R + C_V}{R + 2C_V}.$$ 

Therefore,

$$\Delta S = C_V \log \frac{R + 2C_V}{2(R + C_V)} + R \log \frac{R + C_V}{R + 2C_V}.$$ 

Q10-2. There is a refrigerator that uses external heating process (such as used in camping cars with the use of LPG). Let us imagine an ideal fridge (i.e., reversible fridge) importing heat $Q_H$ from the high temperature heat reservoir (say, a burner) at temperature $T_H$.

Let $T_M$ be the temperature of the campsite. The temperature inside the cooled box is $T_L$ ($T_H \gg T_M > T_L$ is the usual case). The energy balance of the device may be depicted as in Fig. 10.11. For this device to work as a fridge, $Q_H$ and $Q_L$ must be positive (i.e., the device absorbs these heats) and $Q_M$ must be negative (this heat...
Figure 10.11: An idealized LPG fridge.

must be discarded). Since $Q_H$ is supplied by some energy source, the ‘goodness’ of the fridge may be measured by the cost-performance ratio:

$$\eta = \frac{Q_L}{Q_H}. \quad (10.44)$$

(1) Write down the energy balance equation (i.e., $\Delta E = 0$ for a cycle). We strictly apply our sign convention: in $+$, out $-$.

(2) Write down the reversibility condition (i.e., $\Delta S = 0$).

(3) Using these equations, obtain $\eta$ in terms of $T_H$, $T_M$ and $T_L$.

(4) If you look at the obtained $\eta$, you will realize that this ‘goodness measure’ improves (increases) as $T_H$ is raised: hotter the burner, cooler the box! Isn’t it counterintuitive? Explain very briefly why it is not counterintuitive. A hint is the following ‘dissection’ of the fridge in Fig. 10.12 (I do not mean every such fridge contains an engine. The dissection is a conceptual dissection.)

Figure 10.12: Conceptual dissection of the LPG fridge.

**Soln.**

(1) I recommend you to stick to the algebraic sign convention: $Q_H + Q_M + Q_L = 0$.

(2) Since reversibility may be assumed for the ideal case, we may use $dS = d'Q/T$:

$$\frac{Q_H}{T_H} + \frac{Q_M}{T_M} + \frac{Q_L}{T_L} = 0$$
(3) Getting rid of $Q_M$, we get

$$\frac{Q_H}{T_H} + \frac{Q_L}{T_L} = \frac{Q_L + Q_H}{T_M},$$

so

$$\frac{1}{T_H} + \eta \frac{1}{T_L} = (1 + \eta) \frac{1}{T_M},$$

or (recall $T_L < T_M < T_H$)

$$\left(\frac{1}{T_L} - \frac{1}{T_M}\right) \eta = \frac{1}{T_M} - \frac{1}{T_H}.$$

Therefore,

$$\eta = \frac{T_M - T_H}{T_L - T_M} = \left(1 - \frac{T_M}{T_H}\right) \frac{T_L}{T_M - T_L}.$$

This is just the product of the reversible heat engine efficiency working between $T_H$ and $T_M$ and the reversible refrigerator efficiency working between $T_M$ and $T_L$. Thus, Fig. 10.12 is quite natural. This answers (4) as well. Increasing $T_H$ makes the engine efficiency better, so the overall efficiency increases. Thus, the hotter the burner, the cooler the fridge (although we are not actually lowering $T_L$ in this problem).

**Q10-3.** A vertical cylinder of cross section $A$ containing an ideal gas is equipped with a piston and is placed in a room at temperature $T$. Initially, on the piston is a weight of mass $M$ (ignore the ambient pressure, or we do this experiment in the vacuum as illustrated in Fig. 10.13). Now we put another identical weight on the piston. The cylinder is rigid but does not isolate the content thermally. What is the percentage of the potential energy of the weights lost as heat, etc., to the environment?

**Soln.**

Let $V$ be the initial volume. The initial pressure is $Mg/A = P$. The piston moves by $V/2A$, so $W = 2Mg(V/2A) = nRT$ is the potential energy lost from the weights between the initial and the final states. The increase of the free energy of the gas is (notice $\Delta E = 0$ for isothermal process for an ideal gas)

$$\Delta A = \Delta E - T\Delta S = -T\Delta S = -TnR\log \frac{V/2}{V} = nRT \log 2$$

Hence, $100 - 69.3 = 31\%$. We can directly obtain $\Delta A$ as well since $dA = -SdT - PdV$. $T$ is constant, so

$$\Delta A = -\int_{V}^{V/2} \frac{nRT}{V} dV = nRT \log \frac{V}{V/2} = nRT \log 2. \quad (10.45)$$

159
A more detailed explanation is in Fig. 10.14. In this figure, A is the situation we are discussing. The work $W$ done by the weights is the loss of their potential energy $W = nRT$ according to our calculation above. This process is not a gentle process. Let us do this process gently by applying an appropriate force $F$ (B in the figure). Then, the work $W_{\text{rev}}$ (reversible work) done to the gas is the reversible work, so $\Delta A = W_{\text{rev}}$. $W_{\text{rev}}$ is the potential energy difference – the work you did through $F$, so clearly $W > W_{\text{rev}} = \Delta A$. Without your assistance, it is clear that the potential energy of the weights is lost as heat (and perhaps sound), and the loss should be $W - \Delta A$. 

Figure 10.14: A: The actual irreversible process; Left: just after starting the dropping process; Right: the final state. B: By adjusting the force $F$, we wish to lower the weights quasistatically.
11 More examples of $\Delta S$

Summary
* If $\Delta S = 0$, try to devise adiabatic reversible processes.
* $1 \text{ J/K-mol} = 0.173 \text{ bits/molecule}$.
* Helmholtz' free energy gives us the reversible work under isothermal condition. Generally, $\Delta A \leq W$ (pay attention to our sign convention).
* Legendre transformation has a deep meaning.

Key words
entropy of mixing, Helmholtz free energy, Legendre transformation, the Gibbs free energy, enthalpy, free-energy minimum principle

What you should be able to do
* To estimate the entropy change due to various irreversible processes.
* Intuitively understand entropy change in terms of the number of Yes-No questions. $\Delta S > 0$ implies molecules have more freedom (if crudely put), so to determine their states, we must ask more question.
* Draw the general $E = E(S)$ curve under constant work coordinates.

To understand thermodynamics you must memorize the key principles and basic definitions precisely. Recite the key concepts: thermodynamic coordinates/space; extensive vs intensive variables and conjugate pairs; Gibbs relation for quasistatic changes. $\Delta S$ can never be negative, if there is no cooling process. Irreversibility increases $S$ more than corresponding reversible processes (Clausius' inequality). $S$ max condition and equilibrium condition are equivalent, if adiabatic (max entropy principle).

Let us continue thermodynamics through examples.

11.1 If $\Delta S = 0$, there ought to be a reversible process
Suppose, there are two identical containers A and B containing identical amount of water, but their temperatures are distinct: A is at $T_H$ and B at $T_L$ ($< T_H$). A different state that A is at $T_L$ and B at $T_H$ obviously has the same entropy as the former
case. There must be a way to change the initial state to the second one reversibly, i.e., there must be a reversible way to exchange the temperatures only. Can you do this only with heat transfer without using engines? You can use this device to utilize the thermal energy in the used water in a bath tub to warm up the shower water, which may be the usual tap water.

11.2 Sudden doubling of volume
We know adiabatic free expansion is irreversible. Let us double the volume of an ideal gas from $V_I = V$ to $V_F = 2V$ by adiabatic sudden expansion (say, by removing the wall in Fig. 11.1). If the gas is an ideal gas, the total kinetic energy is conserved, so the internal energy of the gas cannot change. This is an irreversible process. Since the process does not change the internal energy of the gas, we can compute the entropy change, using the formula we derived last time, or integrating $dS = (P/T)dV$; this latter approach is equivalent to devising an appropriate quasistatic process. For one mole of the gas

$$\Delta S = R \log \frac{V_F}{V_I} = R \log 2. \quad (11.1)$$

$R = 8.31 \text{ J/K/mol}$, and $\log 2 = 0.693$, so $\Delta S = 5.75 \text{ J/K/mol}$.

Figure 11.1: If the volume is doubled, to locate a molecule as accurately as before expansion we need to know which $V$ (left half or right half) it is in.

11.3 Information: sneak preview
Entropy change and gain or loss of ‘information’ are closely related as we will see in Section 15. Here is a sneak preview. If the volume is doubled, to locate a molecule as accurately as before the expansion we need to know which $V$ (left or right) it is in before knowing at what location inside of one of $V$s the particle is. This knowledge is obtained from a question answered by a single Yes or No (“is it in the left box?”, a single YN question), so the expansion makes a state that requires one bit per molecule\(^{169}\) extra information to describe it relative to the original state (molecules

\(^{169}\)‘bit’ is the unit of information we can obtain from an answer of one yes-no question (YN

162
have acquired more freedom, obviously). Therefore, it is sensible to conclude that the entropy change 5.75 J/K-mol (1 J/K-mol) is interpreted (converted) to the information of 1 bit/molecule (0.174 bit/molecule). 1 J/K-mol entropy increase corresponds to 0.174 bit/molecule of information. Notice that it is per molecule (not per mole); we are asking questions about each molecule.

11.4 What happens if we expand the gas from $V$ to $2V$ reversibly?
This is nothing but adiabatic reversible expansion we have already discussed, but let us look at it from a slightly different angle. $\Delta S = 0$, so the internal energy must be reduced. That is, the gas can do work. You can devise a quasistatic path to $V_F = 2V$. Or, we can consider the final state $(2V, E_F)$ in the thermodynamic space, and demand the process from $(V, E_I)$ to this final state to satisfy $\Delta S = 0$, using the equation of state $S = S(V, E)$ we already know.

First, let us devise a quasistatic process. We must gently expand the gas, so we must apply a force counterbalancing the pressure of the gas. The work done by the gas during this process is

$$\Delta W = -\int_V^{2V} P(V') \, dV'. \quad (11.2)$$

Although $PV = RT$ may be used, the temperature would change. Notice that this is a reversible adiabatic process, so we may use Poisson’s relation $PV^\gamma = \text{constant} = P_I V_I^\gamma$. Therefore,

$$\Delta W = -P_I \int_V^{2V} \left( \frac{V}{V'} \right)^\gamma \, dV' = -P_I V_I^\gamma \frac{1}{1 - \gamma} \left[ (2V)^{1-\gamma} - V_I^{1-\gamma} \right] \quad (11.3)$$

$$= -P_I V_I^\gamma \frac{1}{\gamma - 1} (1 - 1/2^{1-\gamma}). \quad (11.4)$$

This is certainly negative.

Now, let us use the equation of state $S = S(E, V)$. We know $S(E_I, V) = S(E_F, 2V)$, so

$$C_V \log \frac{E_F}{E_I} + R \log \frac{V_F}{V_I} = 0. \quad (11.5)$$

Since $R/C_V = \gamma - 1$, this reads

$$\log \frac{E_F}{E_I} + \log 2^{1-\gamma} = 0 \Rightarrow E_F = E_I/2^{1-\gamma}. \quad (11.6)$$

(question). We will discuss this after the introduction of canonical distribution. Here, simply accept this intuitively.
Indeed, $E_F < E_I$. Can you show $\Delta W = E_F - E_I$ is identical to (11.4)?

### 11.5 Mixing entropy

Mixing of two different substances (even without any interactions between them as in the case of ideal gases) also causes an increase of entropy (Thomson ignored this aspect of entropy). Suppose there are two kinds of ideal gas A and B ($N$ particles each in a separate container of the same volume $V$). They are at the same $T$ and $P$, and can be mixed at constant $T$ and $P$ (due to Dalton’s law of partial pressures) by removing the separating wall at the midpoint of the box (see Fig. 11.2).

![Figure 11.2: The mixing process may be considered as two irreversible volume doublings and subsequent superposition of the expanded gases; the last superposition step does not cause any thermodynamic change, because these gas particles do not interact.](image)

If we use the information-entropy relation above, we can guess the mixing entropy. After mixing, if you pick up a single molecule, you must know whether it is A or B. Before mixing, this information was given ‘for free’, if you know the particle position. That is, mixing process prepares a state that requires one more bit (one extra YN question, say, “is it A?”) to specify the state of its individual molecules. Therefore, $\Delta S = 2Nk_B\log 2$ is our guess (there are $2N$ particles).

The mixing process may be decomposed into the processes illustrated in Fig. 11.2. First, we expand each gas separately to prepare the state at temperature $T$ and pressure $P/2$ (you can do so by adiabatic ‘free’ expansion as discussed just above), and then superpose these two gases to make the final mixture;\(^\text{170}\) since they are ideal gases, they do not feel each other (recall Dalton’s law of partial pressures). Therefore the final superposition step does not cause any thermodynamic change. Therefore, $\Delta S$ must be just the sum of the ‘$V \rightarrow 2V$’ expansion entropy changes, and our guess is correct.

---

\(^{170}\)This is realizable with the aid of semipermeable membranes (walls that can allow only A or B to go through).
11.6 A more general case
A more general case is that the amount of A and B are different; $T$ and $P$ are the same but the volumes are $V_A$ and $V_B$, respectively. Then, the final volume is $V_A + V_B$, so the free expansion entropies for A and B are

$$\Delta S_A = N_A k_B \log \frac{V_A + V_B}{V_A}, \quad \Delta S_B = N_B k_B \log \frac{V_A + V_B}{V_B}. \quad (11.7)$$

That is, the mixing entropy is given by (notice that $P, T$ constant $\Rightarrow N \propto V$)

$$\Delta S = N_A k_B \log \frac{N_A + N_B}{N_A} + N_B k_B \log \frac{N_A + N_B}{N_B}. \quad (11.8)$$

If we introduce the mole fraction $x_A = N_A/(N_A + N_B)$ and $x_B = N_B/(N_A + N_B)$, we can rewrite as

$$\Delta S = (N_A + N_B)k_B(-x_A \log x_A - x_B \log x_B). \quad (11.9)$$

We have learned that if we mix distinct gases A and B, entropy increases. Suppose you have two gases C and D, and wish to know whether they are distinct gases or not. OK, let us measure the mixing entropy. Is this feasible?

11.7 Relaxing isolation/adiabatic conditions
In reality, the variables $S, V, X, \cdots$ for the ordinary Gibbs relation (9.10) are often hard to control or at least awkward. For example, to keep volume constant may be more difficult than to keep pressure constant. To keep the temperature constant may be easier than an adiabatic condition.

11.8 Isothermal system
Under $T$ constant (an isothermal condition) we should allow ‘free’ exchange of heat between the system and its ambient world to maintain the system temperature. Therefore, we wish to pay attention to the RHS of

$$dE - d'Q = d'W = -PdV + xdX. \quad (11.10)$$

Since (11.10) holds under a quasistatic condition, $d'Q = TdS$. $T$ is constant, so (11.10) reads

$$dE - TdS = d(E - TS) = -PdV + xdX. \quad (11.11)$$
This implies the introduction of the quantity
\[ A = E - TS, \] (11.12)
called the Helmholtz free energy.\(^\text{171}\) Notice that for an isothermal process
\[ dA = d'W. \] (11.13)
Thus, \(\Delta A\) is the work the system obtains by a reversible process under constant temperature (i.e., a reversible isothermal process).

### 11.9 \(\Delta A\) by an irreversible process

Work \(W\) is always measurable with the aid of mechanics. What happens if the work exchange is not reversible under isothermal conditions?\(^\text{172}\)

If we inject work \(W\) into the system irreversibly (= that allows some dissipation of work), the system must discard heat to the heat reservoir to maintain its temperature of the final equilibrium state. This implies that the actual changes of the work coordinates \(X\) are less than those due to the same process under a reversible gentle condition. Therefore, even if you do actual work of \(W\), effectively the system receives less energy as work. Therefore, we conclude
\[ \Delta A \leq W. \] (11.14)
Pay attention to the sign convention: ‘coming in is +’! Therefore, (11.14) implies that the work the system can do cannot be larger than \(|\Delta A|\). Suppose the system does work of amount \(|W|\) \((W < 0)\) to the outside. This implies that the system is supplied with the work of \(-|W| = W\), so according to (11.14) \(\Delta A < -|W|\) must hold. Since \(\Delta A < 0\), and \(|\Delta A|\) is the amount of decrease of the system free energy, when the system does work to outside, (11.14) implies
\[ |\Delta A| \geq |W|. \] (11.15)
That is, the work produced by the system cannot exceed the amount of the free energy lost by the system. The work we can gain from the system is bounded by \(|\Delta A|\).

\(^{171}\)Old literatures use \(F\).

\(^{172}\)Strictly speaking, temperature is not definable if a system is not in equilibrium, so you may well question what an isothermal irreversible process means. It means that the initial and the final temperatures are the same. Anything can happen in between. You can also understand the process as occurring in a system immersed in an isothermal bath (thermostat); still there is no guarantee that the system temperature is always well defined.
11.10 Relation to Clausius’ inequality

You might have felt that the above argument sounds like a hand-waving argument (actually, it is not), so let us derive (11.14) from Clausius’ inequality

\[ \Delta S_I \geq \frac{Q}{T}. \]  

(11.16)

Here, we assume \( T_e = T \). \( Q \) is the heat given to system I, so the heat bath loses \( Q \) or gains \(-Q\). Let us assume system II also does work \( W \) to system I. That is, system II gains \(-W\) (however, there is no guarantee that this work is completely received by system I as work). The first law applied to heat bath II reads

\[ \Delta E_{II} = -W - Q. \]  

(11.17)

The definition of the Helmholtz free energy and an isothermal condition imply

\[ \Delta E_I = \Delta A_I + T \Delta S_I. \]  

(11.18)

Since the total energy is conserved (isolation),

\[ 0 = \Delta E = \Delta A_I + T \Delta S_I - W - Q. \]  

(11.19)

Clausius’ inequality implies \( T \Delta S_I - Q \geq 0 \), so this implies

\[ \Delta A_I - W = Q - T \Delta S_I \leq 0. \]  

(11.20)

This is what we wished to have.

11.11 Free energy minimum principle

If there is no exchange of work, irreversibility under isothermal condition implies

\[ \delta A \leq 0. \]  

(11.21)

This implies that, if there is no spontaneous change (i.e., the state is stable), then

\[ \delta A > 0. \]  

(11.22)

That is, in the stable equilibrium state under constant \( T, V, \cdots \) the Helmholtz free energy must be the global minimum. This is the free energy minimum principle.

11.12 Gibbs relation for \( A \)

The Gibbs relation now reads

\[ dA = -SdT - PdV + xdx, \]  

(11.23)

so we see, as designed, the natural set of independent thermodynamic variables is \( (T, V, X) \) instead of \( (S, V, X) \).
11.13 Legendre transformation
Formally, we can say that $E \rightarrow A = E - TS$ allows us to change the independent variables from $(S, V, X)$ to $(T, V, X)$. This is called (in most introductory textbooks) a Legendre transformation. This is, probably, one of the most mysterious parts of thermodynamics, because usually instructors do not know the true meaning of this transformation. Look at the next unit and Appendix 11A to have a real understanding of Legendre transformation.

11.14 Geometrical meaning of Legendre transformation
The transformation $E \rightarrow A = E - TS$ assumes that $T$ in this relation satisfies $\frac{\partial E}{\partial S} = T$. Then, to obtain $A$ is, as can be seen from Fig. 11.3, given $T$, to find on the curve $E = E(S)$ a point where the tangent is $T$ by moving $S$. The difference between the curve $E = E(S)$ and the line $E = TS$ is $A$; we are actually computing $A = \min_S \{E - TS\}$, because to draw a tangent whose slope is $T$ for curve $E = E(S)$ is to minimize the difference between the curve and the line $E = ST$.

![Diagram](image)

Figure 11.3: If we fix $X$'s, $E$ is a monotone increasing convex function of $S$.

As can be seen in Appendix 10A, $-A = \max_S \{TS - E\}$ is mathematically and aesthetically far better, because $E = E(S)$ is a convex function. Then, since Legendre transformation applied twice recovers the original object, $E = \max_T \{ST - (-A)\}$ =

---

\textsuperscript{173}It is shocking that even an expository article of Legendre transformation in Am. J. Phys. does not mention this at all.

\textsuperscript{174}In mathematics, \textit{convexity} always implies downward complexity.
max\(_T\{TS + A\}\) is concluded; we can recover \(E\) from \(A\). Observe the perfect symmetry. [If you use the standard shorthand notation for Legendre transformation, \(-A = E^*\) and \(E = (-A)^*\), i.e., \(E^{**} = E\).]

### 11.15 Gibbs free energy and enthalpy

It is often more convenient to study systems not only under constant temperature but also under constant pressure. Now, the system is placed in a constant pressure thermostat. Then, the work due to the volume change (the volume work \(-PdV\)) must be freely exchanged between the system and the external world, so we should rewrite the Gibbs relation as

\[
dE - TdS + PdV = xdX + \cdots, \quad (11.24)
\]

but since \(T\) and \(P\) are constant, it is convenient to introduce the following *Gibbs free energy* \(G\)

\[
G = E - TS + PV \quad (11.25)
\]

Quite an analogous argument as the case of the Helmholtz free energy tells us that under constant \(T\) and \(P\), if no work other than due to volume changes exists, then

\[
\delta G < 0, \iff \text{spontaneous changes can occur}, \quad (11.26)
\]

\[
\delta G > 0, \iff \text{the equilibrium is stable}. \quad (11.27)
\]

Again, this is the principle of minimum free energy.

The Gibbs free energy may be written as

\[
G = H - TS, \quad (11.28)
\]

where

\[
H = E + PV \quad (11.29)
\]

is called the *enthalpy*.

If there is only volume works, then \(d'W = -PdV\), so under constant pressure the first law reads

\[
dH = dE + PdV = d'Q. \quad (11.30)
\]

That is, the increase of enthalpy is the heat absorbed by the system under constant pressure. Thus, for example, if a chemical reaction occurs in a system, then the change of enthalpy is the reaction heat under constant pressure.
11.16 Entropy changes due to phase transition

Another way to change the system entropy is phase transition, e.g., melting or evaporation. When a solid melts, a latent heat $Q_m$ is absorbed at a constant temperature ($=melting~temperature~T_m$), so the system entropy changes by

$$\Delta S_m = \frac{Q_m}{T_m}. \quad (11.31)$$

Notice that $Q_m$ is measured as the enthalpy change of the system. We have a similar formula for boiling:

$$\Delta S_b = \frac{Q_b}{T_b}, \quad (11.32)$$

where $Q_b$ is the latent heat of evaporation (boiling heat) and $T_b$ is the boiling temperature. For water $\Delta S_m = 21.9$ J/K·mol = 3.7 bits/molecule and $\Delta S_b = 109$ J/K·mol = 18 bits/molecules. Can we understand these entropy changes? Upon melting, water molecules can freely orient in the 3D space. If we simply specify the orientation direction by one of the octants, 3 bits/molecules may not be unreasonable. When evaporated, the volume is expanded by about 1300 times, so even specifying where a molecule is requires extra $\log_2 1300 \approx 10$ bits. Therefore, although we cannot quantitatively explain this 18 bits by such a crude idea, still we can partially understand why $\Delta S_b$ is much larger than $\Delta S_m$.

---

175 We will discuss what phase transition is statistical-mechanically in Section 25.
Appendix 11A: Introduction to Legendre transformation

The true essence of the Legendre transformation is: a convex curve can be reconstructed from the totality of its tangent lines (→Fig. 11.4 Left), 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). $E = E(S)$ and $-A = -A(T)$ are both convex curves.

$$y = f(x)$$

Figure 11.4: **Left**: The totality of tangent lines can recover a convex function. **Right**: $l$ 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 $f^*(\alpha) = \max_x[\alpha x - f(x)]$, then $y = \alpha x - f^*(\alpha)$ is the tangent line in the figure. This gives a geometrical meaning of the Legendre transformation $f \to f^*$.

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. 11.4 Right).$^{176}$

$$f^*(\alpha) = \max_x[\alpha x - f(x)].$$

(11.33)

This is the mathematically standard definition of the Legendre transformation $f \to f^*$.

The essence is: $f$ and its Legendre transform $f^*$ have exactly the same information. That is, if we know $A(T, V)$ we can recover $E(S, V)$. No information is lost (in the last lecture this point will be stressed again; even with phase transitions thermodynamic potentials are equivalent). You may use any convenient thermodynamic potential as you wish to do thermodynamics.

If $f$ is convex, then $f^*$ is convex, and $f^{**} = f$. That is, the inverse Legendre transformation may be given by a symmetric procedure $f(x) = \max_\alpha[\alpha x - f^*(\alpha)]$. $^{176}$‘max’ in such formulas are ‘sup’ in mathematics, but do not worry too much.
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 $g$: $f(x) = \int^x g(x')dx'$.

![Diagram](image)

**Figure 11.5:** Illustration of the relation between $f$ and $f^*$ in 1D.

In (a) of Fig. 11.5 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).

From these illustrations it should be obvious that the relation between $f$ and $f^*$ is perfectly symmetric, so $f^*$ is convex, and $f(x) = \max_\alpha [\alpha x - f^*(\alpha)]$, or $f^{**} = f$. 

172
Q11-1.
(1) In two boxes of the same volume $V$ are identical ideal gases with different pressures but at the same temperature $T$ (Fig. 11.6, so the amount of gases must be different; for simplicity, let’s say 3 moles and 1 mole, respectively) initially.

Figure 11.6: Initially, the pressures are different but the temperatures are identical.

What is the entropy change of the system, when the system reaches a new equilibrium state after the separating wall is punctuated (removed) under adiabatic condition?

**Soln.**
We must devise a quasiequilibrium process between the initial and the final equilibrium states, if we wish to apply thermodynamics. The strategy may be illustrated as Fig. 11.7.

Figure 11.7: How to construct a quasistatic process connecting the initial and the final equilibrium states.

The second step does not cause anything thermodynamically interesting (isn’t it amazing mechanically, however?). Thus, the entropy change, if any, is solely due to the first step in Fig. 11.7. The temperature does not change, so the volume change is the only cause of entropy. Therefore,

$$\Delta S = 3R \log \frac{3V/2}{V} + R \log \frac{V/2}{V} = R \log(27/16). \quad (11.34)$$

What is $\Delta A$?

(2) How about the case in which two boxes in (1) contain distinct ideal gases of the same temperature?

**Soln.**
The mixing process may be dissected into the two steps in Fig. 11.8. The second step does not alter thermodynamics.
Therefore,
\[
\Delta S = 3R \log \frac{2V}{V} + R \log \frac{2V}{V} = 4R \log 2.
\] (11.35)

This is much bigger than (1) due to mixing. Can we understand the result from the information point of view?

**Q11.2** Inside a thermally insulated (i.e., adiabatic) empty (i.e., vacuum) box of volume 10\(V\) is a small can of volume \(V\) containing a one mole of an ideal gas at temperature \(T\). Now, the can is punctuated and the gas leaks out into a larger box and eventually reaches a new equilibrium state (see Fig. 11.9).

**Soln.**

(1) No heat nor work is exchanged with the outside world, so \(\Delta E = 0\). Since our gas is ideal, this implies that the temperature (when definable) is invariant.
We may use the equation of state $S = S(E, V)$. Thus,

$$\Delta S = R \log 10.$$ 

Since $R \log 2$ corresponds to 1 bit/molecule, $\log 10/\log 2 = \log 2 10 = 3.32$ bits. Thus, on the average 3.3 questions. [Recall that doubling of the volume increased the number of yes-no questions needed to specify the microscopic state of a particular molecule by one.]

We will discuss information later in more detail, but this question should be answerable, if you understand the volume doubling and volume quadrupling cases (already discussed). In the former case $\Delta S = R \log 2$ for a one mole ideal gas. This increase corresponds to one extra question about a particular molecule: is it in the right half? In the quadrupling case, $\Delta S = R \log 4$. We must ask two extra questions: is it on the right half? Then, subsequently, is it on the upper half? Thus, $\Delta S/R \log 2$ gives you the number of Yes-No questions you must further ask.

What if the hole made by punctuation is extremely small and molecules can go through it only one by one? No change, because $\Delta E = 0$ does not change, and the final volume does not change, so whatever the process is as long as the system is energetically isolated, the result cannot change. However, you may be suspicious. OK, we can actually compute the entropy change along the actual process.

This is a quasi equilibrium process BUT is not reversible. Actually, if we pay attention to a small portion $\delta N'$ of the gas going out from the can into the box, the process is patently an irreversible expansion. Thus, entropy increases. This increase may be computed, and after integrating all these infinitesimal increases $\delta S$ of entropy, we get exactly the same result (as demonstrated below).

The following detailed calculation is not at all recommended, but let us follow the quasistatic process just described. Suppose $N'$ molecules have already leaked out from the can (assume that the leakage is very slow and quasistatic). Then, the can pressure is $P = (N - N')k_BT/V$. Let $\delta N'$ be the further small amount of leak from the can. Before going out of the can, this portion $\delta N'$ occupies the volume

$$\delta V_i = \frac{\delta N'}{N - N'} V$$

(11.36)

in the can. The pressure of the outer box is $P = N'k_BT/9V$, so the volume of the escaping molecules $\delta N'$ is

$$\delta V_f = \frac{\delta N'}{N'} 9V.$$ 

(11.37)

That is, the leaked $\delta N'$ changes its volume from $\delta V_i$ to $\delta V_f$. Therefore, the entropy increase due to this escape is (notice that the amount of molecules going out is

---

177 This means if you ask 33 yes-no questions about 10 molecules, you can get enough information.
\( \delta N' \)

\[
\delta S = k_B \delta N' \log \frac{\delta V_f}{\delta V_i} = k_B \delta N' \log \frac{9(N - N')}{N'}.
\] (11.38)

We should integrate this from \( N \) to \( N/10 \)(= the remaining amount in the can):

\[
\Delta S = \int_{N}^{N/10} dN' \, k_B \log \frac{9(N - N')}{N'} = N k_B \log 10.
\] (11.39)
12 Introduction to statistical mechanics

Summary
* If we have a translation table between mechanical and thermodynamic quantities, we can calculate thermodynamic quantities with the aid of mechanics.
* The table must include not only mechanical (i.e., thermodynamic coordinates) but thermal quantities. The latter is supplied by Boltzmann’s principle: \( S = k_B \log w(E, X) \).
* With very natural observations as to thermodynamics and mechanics, we can understand this principle from \( dS = d'Q/T \) [as Einstein did].

Key words
phase space, microstate (classical and quantum), microcanonical ensemble, microcanonical partition function, Boltzmann’s principle

What you should be able to do
* Tell what thermodynamics can and cannot do.
* Clearly explain the meaning of the quantities appearing in Boltzmann’s principle.
* Be able to use Boltzmann’s principle for simple examples.
* Explain why the conventional justification of statistical mechanics in terms of ergodic theory is totally absurd.

12.1 Power and limitation of thermodynamics
We have learned rudiments of thermodynamics. As you have realized, thermodynamics is very powerful when right inputs are introduced, but it cannot tell you anything specific to a particular system. For example, the equation of state or the functional form of \( S = S(E, V) \) cannot be supplied by thermodynamics; when we computed this, even for an ideal gas, we relied on \( E = C_V T \) and \( PV = RT \), neither of which is thermodynamically obtained.

You must clearly know what thermodynamics can do and what not. Thermodynamics can compute the changes of state functions (state variables) between two equilibrium states irrespective of the actual processes that have happened, IF the equation of state of the system is known. Thermodynamics cannot calculate materials-specific (or system-specific) properties.
12.2 Why statistical mechanics?
We believe that the microscopic world underlies the world we experience daily, and their descriptions in terms of mechanics is much more detailed than what the macroscopic phenomenology can offer. Then, we can have hope that looking at the microscopic details, we may be able to obtain the information thermodynamics needs but cannot provide.

Since the microscopic world is described by mechanics governing numerous particles, we should try to compute thermodynamic quantities in terms of mechanics. To describe a macrosystem in terms of particle mechanics, we must expect that we need numerous variables far more than the dimension of the thermodynamic space (recall scooping out water on the earth by a tablespoon!). Thus, it is a natural guess that we need some statistical means: statistical mechanics.

12.3 What do we really need?
However, as is emphasized repeatedly, the macroscopic world is the world governed by the law of large numbers, so if you know how to get the expectation values, we do not need statistics explicitly. We need only the translation table of thermodynamic quantities in terms of mechanical quantities.

We have learned that the most fundamental description of any equilibrium state is in terms of thermodynamic coordinates \((E, X)\), where \(E\) is the internal energy and \(X\) (collectively) are the work coordinates. We know \(E\) is the system mechanical energy. \(X\) may be the volume \(V\), magnetization \(M\), etc., and can be described in terms of microscopic mechanical variables easily and/or naturally; we only need their expectation values (no distribution needed). Thus, we can write down the translation table for thermodynamic coordinates relatively easily.

However, thermodynamics is ‘thermo’dynamics. Indeed, we have learned that \(S = S(E, X)\) is the fundamental relation we need in order to use thermodynamics. Therefore, the translation table must include \(S\).

The translation table was completed by Boltzmann in the following form, the Boltzmann principle:\(^{178}\)

\[
S = k_B \log w(E, X),
\]

\(^{(12.1)}\)

where \(k_B\) is the Boltzmann constant, and \(w(E, X)\) is the ‘number’ of ‘microscopic states (= microstates)’ compatible with the macrostate \((E, X)\) (henceforth, the same

\(^{178}\) Some authors define entropy by this formula. However, \(S\) is thermodynamically defined, and we know how to measure \(\Delta S\) (entropy change). The definition of entropy not referring to thermodynamics is empty as a physical concept, because we cannot measure it. The correctness of the translation table is only guaranteed by the fact the it gives correct thermodynamics.

178
symbol $w(E,X)$ will be used to denote the collection of microstates compatible with $(E,X)$ as well. To understand this statement precisely, we must clarify what ‘microstate’ means.\footnote{You may understand that a ‘microstate’ corresponds to an ‘elementary event’ in probability.}

![Diagram](https://www.dropbox.com/home/IntroQM/contents_files)

**Figure 12.1:** For each equilibrium state $(E,X)$ in the thermodynamic space, we can imagine a subset $w(E,X)$ of the phase space consisting of microstates that give the same thermodynamic coordinates.

### 12.4 Classical microstates

Classical-mechanically,\footnote{The Feynman Lectures I is a good classical mechanics introduction. Then, read the first volume of the Landau-Lifshitz series.} the most detailed description of a system is in terms of a set of the canonical variables. The most popular canonical variables are the position and momentum vectors. For an $N$-point-mass system, the $6N$-dimensional vector $(r_1, \cdots, r_N, p_1, \cdots, p_N)$, where $r_i$ is the position vector of the $i$th particle and $p_i$ the momentum vector of the $i$th particle, gives the ultimately detailed description of the system. The space spanned by these $6N$ coordinates (the totality of these $6N$-dimensional vectors) is called the *phase space* of the system, and a point in this space is classically the elementary event = microstate.

### 12.5 Quantum microstates

Quantum-mechanically,\footnote{The Feynman lectures III is a good quantum mechanics introduction. However, if you wish to finish the rudiments as quickly as possible, read Griffiths. At a more leisurely pace, if you are interested in a more historical development, see the beginning part (Part I) of my lecture notes \url{https://www.dropbox.com/home/IntroQM/contents_files}.} an elementary state is (roughly speaking) a state ket, but
physically $|\cdot\rangle$ and $c|\cdot\rangle$ for any complex number $c$ are indistinguishable, so actually, an elementary state is a ray ($=1$D subspace spanned by a ket). We may take a convenient orthonormal basis of the vector space spanned by all the state kets and interpret any vector in this basis set as a microstate. In particular, normalized eigenkets of the system Hamiltonian may be interpreted as microstates.\footnote{Any unitary transformation of the basis set is again a basis set, so in quantum mechanics the choice of the microstates is not unique. The situation is not different for classical cases; we may apply any canonical transformation to the phase space coordinates.}

### 12.6 How to obtain microcanonical partition function

Thus, to obtain $w(E, X)$, for classical cases, we calculate the volume of the subset of the phase space compatible with the thermodynamic coordinates $(E, X)$. Recall that for any phase point we can compute $(E, X)$ (at least in principle). If the computed $E'$ and $X'$ nearly agree with (that is, are macroscopically indistinguishable from) the thermodynamic coordinates of a macrostate $(E, X)$, we say that the microstate is compatible with this macrostate (see Fig. 12.1). Thus, we can find the subset $w(E, X)$ of the phase space consisting of such microstates.

Quantum-mechanically, we make observables corresponding to $E$ (that is, the system Hamiltonian) and $X$ (we may write such an operator as $\hat{X}$), and then collect eigenkets $|\cdot\rangle$ of the Hamiltonian whose eigenvalues are close to $E$ and also $\langle \cdot|\hat{X}|\cdot\rangle \simeq X$.

The set $w(E, X)$ is called a microcanonical ensemble, and the numerical value $w(E, X)$ (phase volume or number of states) is called a microcanonical partition function.

### 12.7 Statistical Mechanics is completed!

We have completely specified the statistical mechanical rule to compute thermodynamics. The rest is taken care of by the Gibbs relation

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN - \frac{x}{T}dX + \cdots \tag{12.2}$$

In practice, use statistical mechanics sparingly, and use thermodynamics whenever you can.

Of course, there are two problems remaining: How to use Boltzmann’s principle and how to understand the principle. First, let us use the principle a bit.
12.8 Let us study an ideal gas (classically)

We can use the completed translation table to compute $S$ from mechanics. Let us study a classical ideal gas. It consists of $N$ non-interacting mass points of mass $m$ in a volume $V$. The system Hamiltonian is the pure kinetic energy:

$$ H = \sum_i \frac{p_i^2}{2m}. \quad (12.3) $$

Precisely speaking, $w(E, V)$ collects all the microstates in the volume $V$ with energy in $(E - \Delta E, E]$, where $\Delta E$ is a (macroscopically small) leeway. What we should do first is to formally write down $w(E, V)$.

$$ w(E, V) = \int_{r_i \in V, E - \Delta E < \sum p_i^2/2m \leq E} d\Gamma, \quad (12.4) $$

where $d\Gamma = dr_1 \cdots dr_N dp_1 \cdots dp_N$ is the volume element of the $6N$-dimensional phase space. The space and momentum integrals can be totally decoupled, so

$$ w(E, V) = \int_{V} dr_1 \cdots \int_{V} dr_N \int_{\sum p_i^2/2m \in (E - \Delta E, E]} dp_1 \cdots dp_N \quad (12.5) $$

$$ = V^N \int_{\sum p_i^2/2m \in (E - \Delta E, E]} dp_1 \cdots dp_N. \quad (12.6) $$

The last integral is the volume of the skin of thickness $\propto \Delta E$ of a $3N$ dimensional ball ($3N-1$-sphere\(^{183}\)) of radius $\sqrt{2mE}$, which must be proportional to $E^{(3N-1)/2} \Delta E$,\(^{184}\) so

$$ w(E, V) \propto V^N E^{3N/2} \Delta E. \quad (12.7) $$

Here, $N \gg 1$, so 1 is ignored. Therefore, Boltzmann tells us that

$$ S = k_B \log w(E, V) = Nk_B \log V + \frac{3}{2} Nk_B \log E + k_B \log \Delta E + \cdots, \quad (12.8) $$

where the remaining terms are $N \times$ a constant.\(^{185}\) Using thermodynamic relations,

$$ \frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_V = \frac{3}{2} \frac{Nk_B}{E} \quad (12.10) $$

\(^{183}\)Notice that in mathematics, 1-sphere is the edge of a disk, 2-sphere is the skin of a 3D ball (= the ordinary sphere), etc.

\(^{184}\)Dimensional analysis can give you this answer as well.

\(^{185}\)By the way, $\Delta E$ can be pretty much anything, if not too small (as $\Delta E = O[E/\epsilon^{aN}]$ for some $a > 0$).
or \( E = \frac{3}{2} N k_B T \). Also
\[
\frac{P}{T} = \frac{\partial S}{\partial V} \bigg|_E = \frac{N k_B}{V},
\]
which is the equation of state. It works! Boltzmann also used an ideal gas to check that his ‘hypothesis’ works as you can see from Appendix 11A.

Incidentally, (12.9) does not satisfy the fourth law: if you double extensive quantities (experimentally, we have only to join two identical systems to make a compound system): \( N \to 2N, E \to 2E, V \to 2V \), then we must also have \( S \to 2S \), but this does not hold. When you use statistical mechanics, a wise practice is to make a shortcut with the aid of thermodynamics. We demand that the fourth law holds. Then, we are forced to accept the following form:
\[
S = k_B \log w = N k_B \log \frac{V}{N} + \frac{3}{2} N k_B \log \frac{E}{N} + k_B \log \Delta E + \cdots.
\]
This corresponds to replacing \( w \) with \( w/N! \). Boltzmann also noted in his original paper that the latter choice is convenient, because the extensivity of entropy is satisfied. We will come back to this problem in Lecture 13.

### 12.9 Basis of Boltzmann’s formula: key observations

Boltzmann’s formula works, so don’t ask any question and get good grades and publish papers. Wise professors may well preach like this. Well, science should not tolerate incantations. Every black box must be opened.

Let us go to a derivation(!) of Boltzmann’s principle starting from empirical facts and some general observation about mechanics. Since we are physicists, and not metaphysicists, let us be as empirical as possible.

The facts we wish to rely on are:

- [O] If we (in a constant environment thermally) isolate a system, it will eventually arrive at a system that does not depend on time macroscopically (actually this must be the main part of the zeroth law).
- [X] The needed observation time for thermodynamic coordinates is very short (say, 1 \( \mu s \) or much less if the system is large enough).

A general picture of microscopic dynamics is that the instantaneous microstate wanders around the phase space; in particular, if the macrostate is in \((E, X)\), it wanders around in \( w(E, X) \). Traditionally, ‘statistics’ of ‘statistical mechanics’ was understood as taking statistics over all the microstates in \( w(E, X) \). Consequently, a misconception was spread that the key to statistical mechanics was the even mechanical sampling over \( w(E, X) \) (the ergodic theoretical justification of statistical
mechanics).

For the ordinary macro object containing $N \sim 10^{23}$ particles, what is the time scale required to sample $w$ evenly? It is roughly the time scale of the Poincaré cycle $\sim e^N$.\footnote{The time scale of the Poincaré cycle is roughly the time scale required for a given closed dynamical system to return to its initial condition. Poincaré’s recurrence theorem guarantees that any state can return to its any neighborhood.} [X] implies that during one thermodynamic observation, only an extremely tiny fraction of $w$ is sampled. However, [O] implies that if you repeat this experiment after 1 billion years later, we get the same thermodynamic result. A 1 $\mu$s observation in $10^9$ CE will cover again only very tiny portion of $w$ (Fig. 12.2).

Figure 12.2: Repeating thermodynamic observation samples extremely tiny subsets (black dots) of $w$, but they all give the same thermodynamic results.

What is the most natural conclusion? Suppose we prepare an isolated macrosystem reaching an equilibrium state $(E, X)$. Since it is isolated, we may interpret it as a mechanical system as well. If you sample a microstate (mechanically instantaneous state), and compute the thermodynamic coordinates, almost surely they agree with $(E, X)$ and give the correct thermodynamic relation. That is, if you wish to know thermodynamics, you have only to sample a single microstate; almost every microstate gives the same thermodynamics. This is the secret of equilibrium statistical mechanics.

12.10 Derivation of Boltzmann’s formula
Let us derive Boltzmann’s principle from thermodynamics (and mechanics). For simplicity, only $E$ is written explicitly; you can repeat the following argument with $X$ restored. The phase volume $w(E)$ compatible with the microstates whose energy is in $(E - \Delta E, E]$ may be written as

$$w = \int d\Gamma \chi_{\Delta E}(H - E), \quad (12.13)$$

where $\chi_{\Delta E}$ is the indicator of the leeway set, $(-\Delta E, 0]$, and $d\Gamma$ is the phase
Here, the integration is over all the phase space of the system. Let us assume that the Hamiltonian contains a parameter $\lambda$ that can be controlled externally (to do macroscopic work, or to regulate work coordinates; $\lambda$ is something like a handle). The change of the Hamiltonian due to the change of the parameter $\lambda \rightarrow \lambda + \delta \lambda$ (averaged over the original equilibrium distribution) is identified with work $\delta' W$ by Einstein. We can also change $E$ by $\delta E$. Thus, notice that

$$\langle H(\lambda + \delta \lambda) - H(\lambda) \rangle - \delta E = \delta' W - \delta E,$$

(12.14)

where $\langle \rangle$ is the equilibrium average. In terms of this variation, we have

$$\delta w = \int \chi'_{\Delta E} (H - E) (H(\lambda + \delta \lambda) - H(\lambda) - \delta E) d\Gamma.$$

(12.15)

Since almost all the microstates give the same thermodynamic results as already argued, $\delta' W - \delta E$ is almost always the same for any microstate compatible with a given thermodynamic state. Therefore, we obtain

$$\delta w = (\delta' W - \delta E) \int \chi'_{\Delta E} (H - E) d\Gamma,$$

(12.16)

or with the aid of the first law $\delta' W - \delta E = -\delta' Q$, after dividing with $w$,

$$\delta \log w = -\delta' Q \frac{\int d\Gamma \chi'_{\Delta E} (H - E)}{\int d\Gamma \chi_{\Delta E} (H - E)}.$$

(12.17)

Defining

$$\eta \equiv \left. \frac{\partial \log w}{\partial E} \right|_X = -\frac{\int d\Gamma \chi'_{\Delta E} (H - E)}{\int d\Gamma \chi_{\Delta E} (H - E)},$$

(12.18)

we get (under constant $X$)

$$d \log w = \eta d' Q.$$

(12.19)

What is $\eta$? Let us compute this quantity for an ideal gas. Actually, we have already done that in (12.10):

$$k_B \frac{\partial \log w(E, V)}{\partial E} = \frac{3k_B N}{2E}.$$

(12.20)

We know this is $1/T$ even before thermodynamics was established (recall the Maxwell distribution). Now we can use thermodynamics $\delta S = \delta' Q/T$, (and adjusting the units, if needed), we conclude

$$dS = k_B d \log w.$$

(12.21)

\[187\] A quantum version is in PST 17.12.
Integrating this, we obtain Boltzmann’s principle.  

12.11 Derivation of Boltzmann’s formula: quick way

If we accept that entropy is a functional of \( w \), then (12.21) is an inescapable conclusion.

A crucial observation is that entropy is an extensive quantity. If we form a compound system by combining two systems I and II already in thermal equilibrium with each other, the entropy of the compound system is the sum of that of each component (the fourth law).

The interaction introduced by the contact of the two systems is, for macroscopic systems, a very weak one. In any case, the effect is confined to the boundary layer whose thickness is microscopic. Thus, the two subsystems may be regarded statistically independent. Therefore, the total number of microstates of the compound system must be very close to the product of the total numbers of microstates for I and II: \( w = w_I w_{II} \).

Combining the above considerations, we have arrived at the following functional relation:

\[
S(w_I w_{II}) = S(w_I) + S(w_{II}),
\]

(12.22)

where suffixes denote subsystems.

Assuming that \( S \) is an increasing function of \( w \), we may conclude from the relation that \( S \) is proportional to \( \log w \). Therefore, we have arrived at (12.21). The proportionality coefficient \( k_B \) must be positive, because entropy should be larger with larger \( E \) that corresponds to larger \( w \).

12.12 Entropy and information

We have already discussed the meaning of entropy in terms of the number of yes-no questions to specify the molecular state. From this idea, notice that \( S \propto \log w \) is quite natural. \( S \) in bits is the number of YN questions you must ask to pinpoint the microstate when you know the system is in a particular macrostate.

12.13 Why traditional justification is meaningless

When Boltzmann arrived at his statistical mechanics framework, he initially thought that thermodynamic observables were the average value over \( w \). To average, the trajectory of the system as a mechanical system should sample the subset \( w \) evenly, so

\[185\]
he conceived the so-called ergodicity: the trajectory can visit in any neighborhood of any point in \( w \). Almost all the currently popular textbooks explain this totally wrong idea. You should have already realized this.

As we have seen in Lecture 7, every ambitious young man attempted to derive the irreversibility from mechanics. Boltzmann studied the gas dynamic in detailed and 'demonstrated' irreversibility. His colleague Loschmidt questioned why time-reversal symmetric mechanics could give rise to a system losing this symmetry. Boltzmann realized that there is an approximation to discard memory. Then came Zermelo who pointed a logical error out: even if memory is discarded, still any trajectory can return in any neighborhood of the starting point (Poincaré’s recurrence theorem), so irreversibility cannot be concluded. Boltzmann countered that the young man (= Zermelo) should know physics; can you wait for that long time of order \( 10^N \)? Thus, in practice, irreversibility is real, even if the system is finite.

However, you must have quickly realized that this counterargument backfires. For an even sampling of the phase space, you must observe the system for an extremely long time. That is, ergodicity cannot justify statistical mechanics! Actually, Boltzmann seems to have realized that there was a serious problem with the ergodicity argument to found statistical mechanics long before Zermelo’s criticism. He even realized the secret of equilibrium statistical mechanics: every microstate gives the same thermodynamic observables!

Boltzmann’s followers all ignored (or could not understand) this insight. The total absurdity of the ergodicity argument should be obvious from the fact that larger systems require shorter observation times for accurate determination of thermodynamic observables.
Q12.1. Let us taste the original paper.\textsuperscript{189,190} 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

$$M = \sum w_n \log w_n.$$  \hfill (12.23)

2. Write $w_i = w(x)\varepsilon$ 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

$$M' = \int w(x) \log w(x) dx.$$  \hfill (12.24)

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 $M'$.

4. Now, Boltzmann realized that $\log Z_K = \log N! - M'$ gives the entropy of the ideal gas. Based on this finding, he proposed

$$S \propto \log \text{(Number of ‘Komplexions’).}$$  \hfill (12.25)

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

Remark. Usually, the story ends here (so did Boltzmann’s original paper). However, being a much deeper thinker than is usually regarded, Boltzmann later confirmed for macrosystems described by $E$ and $V$ that his formula of entropy (12.1) satisfied the Gibbs relation for general classical many-body systems; in particular, $(dE + PdV)/T$ is a complete differential.\textsuperscript{191}

Soln.

\textsuperscript{189}This is adopted from \textit{Equilibrium Statistical Mechanics} (ver May 2012).

\textsuperscript{190}L. Boltzmann, “Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Sätzen über Wärmeungleichgewicht,” Wiener Berichte \textbf{76}, 373 (1877) (“On the relation between the second law of thermodynamics and probability calculation concerning theorems of thermal equilibrium”)

\textsuperscript{191}For a detail, see Q17.3 in PST.
(1) The Komplexionszahl reads

\[ Z_K = \frac{N!}{w_1!w_2!\cdots w_n!}, \quad (12.26) \]

so maximizing this is equivalent to minimizing the denominator or its logarithm (here Stirling’s formula (13.6) is used):

\[ \log(w_1!w_2!\cdots w_n!) = \sum_n \log w_n! = \sum_n (w_n \log w_n - w_n) = M - N. \quad (12.27) \]

The original paper kindly discusses that we can discard numerical factors, etc., in Stirling’s formula, which we will derive in the next lecture.

(2) Substituting the quantities in \( M \) as indicated, we have

\[ \sum_n w_n \log w_n = \sum_n w(x) \varepsilon \log[w(x)\varepsilon] = \sum_n w(x) \varepsilon \log w(x) + \sum_n w(x) \varepsilon \log \varepsilon \quad (12.28) \]

The first term is a Riemann sum, so we obtain (12.24). The second term is \( N \log \varepsilon \) and is unrelated to the number of complexions, so we may ignore it.

(3) 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 \( p \) and the position \( r \) as \( x \): \( w(x) = w(r,p) \). The constraints are

\[ N = \int dr dp w(r,p), \quad E = \int dr dp w(r,p) E(p), \quad (12.29) \]

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

\[ \int dr dp w(r,p) [\log w(r,p) + \alpha + \beta E(p)]. \quad (12.30) \]

Hence, \( (E = (3/2)Nk_B T \) is used to fix \( \beta \)

\[ w(r,p) = \frac{N}{V} \frac{1}{(2\pi mk_B T)^{3/2}} e^{-p^2/2k_B T m}. \quad (12.31) \]

Thus, we have obtained the Maxwell distribution.

(4) If we compute (12.25) (i.e., \( N \log N - M' \)) with the aid of \( w \) in (12.31)

\[ S = N \log V + N \left( \frac{3}{2} + \frac{3}{2} \log(2\pi k_B T m) \right) = N \log VT^{3/2} + \text{const.} \quad (12.32) \]
This agrees with the entropy obtained thermodynamically (apart from $\log N!$; the formula is not extensive). Indeed, the Gibbs relation is

$$dS = \frac{1}{T}dE + \frac{P}{T}dV,$$  \hfill (12.33)

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

$$S = Nk_B \log T^{3/2} + Nk_B \log(V/N) + \text{const.}$$  \hfill (12.34)
13 Statistical mechanics of isothermal systems

Summary
* Isothermal systems are handled by the canonical formalism: \( A = -k_B T \log Z \).
* Microcanonical and canonical formalisms give identical results if the system is large (if \( \log N/N \ll 1 \)) [Ensemble equivalence].
* The principle of equal probability allows us to estimate the probabilities of a collection of microstates.

Key words
canonical partition function, microcanonical partition function, Gibbs-Helmholtz formula, ensemble equivalence, Stirling’s formula, Schottky defect, Schottky type specific heat.

What you should be able to do
* You must be able to compute the microcanonical partition functions and canonical partition functions for simple systems.
* You must remember the Gibbs formula for \( S \) (i.e., \( dS = \cdots \)).

13.1 What we have now
We have constructed the statistical mechanics for isolated systems = the translation table between mechanical and thermodynamic quantities: for the thermodynamic coordinates the correspondence is straightforward. To utilize the Gibbs relation

\[
dE = TdS - PdV + \mu dN + xdX + \cdots ,
\]

we need the interpretation of \( S \) in terms of mechanics. We have derived Boltzmann’s principle:

\[
S = k_B \log w(E, X).
\]

\( w \) is not very easy to compute, so we will not discuss how to use this formula very much, but let us look at a simple example.

13.2 Another practice: Schottky defects
Let us consider an isolated crystal with point defects (vacancies) on the lattice sites (Schottky defects). To create one such defect we must move an atom from a lattice.
point to the surface of the crystal. The energy cost for this is assumed to be $\varepsilon$.

Although the number $n$ of vacancies are macroscopic, we may still assume it to be very small compared to the number $N$ of all lattice sites. Hence, we may assume that the volume of the system is constant. Therefore for this example, the (internal) energy $E$ of the system is a macroscopic (thermodynamic) variable which completely specifies macrostates.

We must compute $w$ as a function of the total energy $E$, which is given by

$$E = n\varepsilon.$$  \hfill (13.3)

We may interpret this as the internal energy. We may consider $w$ as a function of $n$. A microstate of this system is specified by the locations to place $n$ vacancies. Since all the lattice points can be distinguished, the number of placing $n$ vacancies is obviously

$$w(n) = \binom{N}{n}. \hfill (13.4)$$

To compute the entropy with the aid of Boltzmann’s principle, we use Stirling’s formula to evaluate $\log N!$ asymptotically for large $N$:

$$N! \approx \left(\frac{N}{e}\right)^N, \hfill (13.5)$$

or

$$\log N! \approx N \log N - N, \hfill (13.6)$$

which may be understood as follows:

$$\log N! = \sum_{k=1}^{N} \log k \simeq \int_{0}^{N} dx \log x = (x \log x - x)_{x=0}^{N} = N \log N - N. \hfill (13.7)$$

Boltzmann’s principle gives us

$$S = k_B \log w \simeq k_B [N \log N - n \log n - (N - n) \log (N - n)]. \hfill (13.8)$$

Incidentally, the following formula is useful (and easy to remember):

$$\log \left(\frac{A}{B}\right) = -A \left[\frac{B}{A} \log \frac{B}{A} + \left(1 - \frac{B}{A}\right) \log \left(1 - \frac{B}{A}\right)\right]. \hfill (13.9)$$

This gives

$$S = -nk_B \left[\frac{n}{N} \log \frac{n}{N} + \left(1 - \frac{n}{N}\right) \log \left(1 - \frac{n}{N}\right)\right], \hfill (13.10)$$
which is the same as (13.8). Using the Gibbs relation, we get (notice that $dE = \varepsilon dn$)

$$
\frac{1}{T} = \frac{\partial S}{\partial E}_{V} = \frac{1}{\varepsilon} \frac{dS}{dn} = \frac{k_B}{\varepsilon} \log \frac{N - n}{n}.
$$

(13.11)

When you differentiate (13.10), the derivatives of the logarithmic terms cancel each other, so essentially, you have only to differentiate the prefactors in front of the logarithms. This is very easy, and you immediately get (13.11).

If the temperature is sufficiently low or $\varepsilon$ is sufficiently large so that $\varepsilon/k_B T \gg 1$, the above formula reduces to

$$
\frac{\varepsilon}{k_B T} \simeq \log \frac{N}{n},
$$

(13.12)

because $N \gg n$. Hence, under this low temperature condition, the internal energy $E$ reads

$$
E = \varepsilon N e^{-\varepsilon/k_B T},
$$

(13.13)

which may be guessed from the Boltzmann factor.

### 13.3 Schottky type specific heat

The constant volume specific heat $C_V$ of the system can be obtained as

$$
C_V = \frac{dE}{dT} = N k_B \left( \frac{\varepsilon}{k_B T} \right)^2 e^{-\varepsilon/k_B T}.
$$

(13.14)

Notice that $C_V$ has a peak at a certain temperature (Fig. 13.1). This type of specific heat is called the *Schottky type specific heat*, which tells you the energy gap for an elementary excitation in the system. What can you say from the dimensional analytical point of view about the same problem?

![Figure 13.1: The Schottky type specific heat, which has a peak indicating the energy gap of the order $k_B T_P = \varepsilon/2$.](image)
13.4 Thermostat

Isolated systems are not so easy to handle, compared with thermostatted systems. We have introduced the Helmholtz free energy $A$ to study thermostatted systems. We have learned that the information as to thermodynamic potentials can be completely obtained from $A$ (the equivalence of thermodynamic potentials thanks to the Legendre transformation).

Let us extend the formalism for isolated systems to thermostatted systems. The logic we use is the familiar one: we regard the system as a small portion I of a big isolated system, and assume that the system can freely exchange heat with its surroundings II.

13.5 Statistical mechanics of thermostatted system: heuristics\textsuperscript{192}

The total energy $E_0$ of the compound system is given by

$$E_0 = E_I + E_{II}. \quad (13.15)$$

The number of microstates for system I (resp., II) with energy $E_I$ (resp., $E_{II}$) is denoted by $w_I(E_I)$ (resp., $w_{II}(E_{II})$). Thermal contact is a very weak interaction, so the two systems are statistically independent. Hence, the number of microstates for the compound system with the energies $E_I$ in I and $E_{II}$ in II is given by

$$w_I(E_I)w_{II}(E_{II}). \quad (13.16)$$

The total number $w(E_0)$ of microstates for the compound system must be the sum of this product over all the ways to partition energy between I and II. Therefore, we get

$$w_{I+II}(E_0) = \sum_{0 \leq E_I \leq E_0} w_I(E_I)w_{II}(E_0 - E_I). \quad (13.17)$$

The system II is huge compared with I. Expand the entropy as follows:

$$S_{II}(E_0 - \mathcal{E}) = S_{II}(E_0) - \mathcal{E} \frac{\partial S_{II}}{\partial E_{II}} + \frac{1}{2} \mathcal{E}^2 \frac{\partial^2 S_{II}}{\partial E_{II}^2} + \cdots \quad (13.18)$$

\textsuperscript{192}‘Difficulty of standard explanation’ What you will read here is a more or less standard conventional explanation. Unfortunately, this argument has two flaws.

(1) The twice differentiability of $S$ as a function of $E$ is required to justify the use of the Taylor expansion, but this excludes many phase transitions.

(2) The outer thermostat is used in order to save the description of the whole system as an isolated mechanical system, but the larger the system is, the harder the pure mechanics becomes. Thus, we need a completely different justification (see 14.4).
and denote the temperature of the heat bath (i.e., system II) by $T$:

$$\frac{\partial S_{\text{II}}}{\partial E_{\text{II}}} = \frac{1}{T}. \quad (13.19)$$

We wish to use this formula in equilibrium, so $\mathcal{E}$ should be close to the internal energy of system I. Therefore, due to the extensivity of internal energy this should be of order $N_1$, the total number of particles in system I. Therefore,

$$\mathcal{E} \frac{\partial S_{\text{II}}}{\partial E_{\text{II}}} = O[N_1]. \quad (13.20)$$

The second derivative in (13.18) is proportional to $1/\frac{\partial E}{\partial T} = 1/C_{V\text{II}}$, where $C_{V\text{II}}$ is the specific heat of II, which is $O[N_2]$:

$$\mathcal{E}^2 \frac{\partial^2 S_{\text{II}}}{\partial E_{\text{II}}^2} = -\frac{\mathcal{E}^2}{T^2 C_{V\text{II}}} = O[N_1]^2 \frac{O[N_1]}{O[N_2]} \ll O[N_1]. \quad (13.21)$$

Therefore, the ratio of the second term and the third term in (13.18) is of order $N_1/N_2$, which is negligibly small. Thus, (13.17) reads

$$w_{I+\text{II}}(E_0) \simeq e^{S_{\text{I}}(E_0)/k_B} \sum_{\mathcal{E}} w_1(\mathcal{E}) e^{-\beta \mathcal{E}}, \quad (13.22)$$

or

$$w_{I+\text{II}}(E_0) e^{-S_{\text{II}}(E_0)/k_B} \simeq \sum_{\mathcal{E}} w_1(\mathcal{E}) e^{-\beta \mathcal{E}}, \quad (13.23)$$

where a standard notation

$$\beta = 1/k_B T \quad (13.24)$$

is used.

With the aid of Boltzmann’s principle, we have, with $E$ being the equilibrium internal energy of system I,

$$k_B \log w_{I+\text{II}}(E_0) = S_I(E) + S_{\text{II}}(E_0 - E), \quad (13.25)$$

so (from now on let’s drop the suffix I to denote the system)

$$k_B \log [w_{I+\text{II}}(E_0) e^{-S_{\text{II}}(E_0)/k_B}] = S(E) + S_{\text{II}}(E_0 - E) - S_{\text{II}}(E_0) = S(E) - E/T = -A/T. \quad (13.26)$$

That is, (13.23) reads (suffix I dropped from $w_I$)

$$e^{-\beta A} = \sum_{\mathcal{E}} w(\mathcal{E}) e^{-\beta \mathcal{E}}. \quad (13.27)$$
13.6 Canonical formalism

Thus, we have arrived at our desired formalism, the canonical formalism: Let

\[ Z = \sum_{\mathcal{E}} w(\mathcal{E}) e^{-\beta \mathcal{E}}. \tag{13.28} \]

Then,

\[ A = -k_B T \log Z. \tag{13.29} \]

\(Z\) is called the canonical partition function, and this method to compute the free energy is called the canonical formalism.

A more microscopic expression is possible:

\[ Z = \sum_{\text{all microstates}} e^{-\beta \mathcal{E}} = \text{Tr} e^{-\beta \mathcal{H}}. \tag{13.30} \]

Here, the sum over all the microstates is, in the quantum mechanical cases, the summation over all the eigenvalues of the Hamiltonian, so quantum mechanically, we may use the trace to compute the partition function. If we decompose the sum as follows, we can easily understand this formula:

\[ \sum_{\text{all microstates}} = \sum_{\mathcal{E}} \sum_{\text{all microstates with energy } \sim \mathcal{E}}, \tag{13.31} \]

but

\[ \sum_{\text{all microstates with energy } \sim \mathcal{E}} e^{-\beta \mathcal{E}} = w(\mathcal{E}) e^{-\beta \mathcal{E}}. \tag{13.32} \]

13.7 Principle of equal probability

The conventional more or less standard statistical mechanics assumes a principle called the principle of equal probability: if we sample a microstate from \(w(E, X)\) every microstate is equally probable. Why this principle may be used to understand thermodynamic quantities is explained in Appendix 13A. If we accept this principle, then in each \(w(E, X)\) the probability to sample any subset \(u \in w(E, X)\) is proportional to its phase volume (classically) or number of states in it (quantum-mechanically). Thus, we can interpret that the probability for a microstate \(\gamma\) is

\[ P(\gamma) = \frac{1}{Z} e^{-\beta H(\gamma)}. \tag{13.33} \]
This is called the *canonical distribution*.

### 13.8 Warning about the canonical distribution

(13.33) may look as if it gives the probability for individual microstates. However, for a macroscopic system there is no way to single out individual microstates, so there is no direct way to verify (13.33) experimentally. Thus, the expression itself is only formal and no physical meaning as it is. A thermodynamic justification of the principle of equal probability (see Appendix 13A) corroborates this warning.

(13.33) is legitimate only when we compute expectation values of macro or mesoscopic observables or probabilities of macro or mesoscopic events.

### 13.9 The Gibbs-Helmholtz formula

Once the canonical partition function is known, the internal energy of the system can be obtained easily:

\[
E = \langle \mathcal{E} \rangle = \sum_{\mathcal{E}} P(\mathcal{E}) \mathcal{E} = \frac{1}{Z} \sum_{\mathcal{E}} \mathcal{E} w_1(\mathcal{E}) e^{-\beta \mathcal{E}} = -\frac{\partial \log Z(\beta)}{\partial \beta},
\]

where \( Z \) (cf. (13.30)) is explicitly written as a function of \( \beta \). (13.34) is a thermodynamically well-known formula:

\[
\frac{\partial (A/T)}{\partial (1/T)} \bigg|_V = E, \quad \text{or} \quad \frac{\partial \beta A}{\partial \beta} \bigg|_V = E,
\]

the *Gibbs-Helmholtz formula*. Do not forget that this is a purely thermodynamic relation.

### 13.10 Schottky defects revisited

Let’s revisit the Schottky defects. With \( w(n) \) known (see (13.4)), it is easy to compute \( Z \):

\[
Z = \sum_n w(n) e^{-\beta n \varepsilon} = (1 + e^{-\beta \varepsilon})^N,
\]

where we have used the *binomial theorem*:

\[
(x + y)^N = \sum_{n=0}^{N} \binom{N}{n} x^n y^{N-n}.
\]
If you are uncomfortable, review Appendix 2A after Lecture 2. Thus,

\[ \sum_{n=0}^{N} w(n) e^{-\beta n \varepsilon} = \sum_{n} \binom{N}{n} (e^{-\beta \varepsilon})^n 1^{N-n} = (1 + e^{-\beta \varepsilon})^N. \] (13.38)

However, you can probably write down the right-most formula directly: the canonical partition function is a sum over all the possible microstates

\[ Z = \sum_{\varepsilon(1)\in\{0,\varepsilon\},\cdots,\varepsilon(N)\in\{0,\varepsilon\}} e^{-\beta \sum_{i=1}^{N} \varepsilon(i)}, \] (13.39)

where \( \varepsilon(i) \) is the energy of the \( i \)th lattice point (occupied 0 or empty \( \varepsilon \)). Here, do not forget that a ‘microstate’ is a microscopically described state of the whole macro system; in our case \( (\varepsilon(1), \varepsilon(2), \cdots, \varepsilon(N)) \) is a microstate, where each \( \varepsilon(i) \) is 0 or \( \varepsilon \). Do not confuse the microstate and the elementary states of individual microscopic entities. Notice that all the combinations of the lattice states show up, so

\[ Z = \left( \sum_{\varepsilon(1)\in\{0,\varepsilon\}} e^{-\beta \varepsilon(1)} \right) \cdots \left( \sum_{\varepsilon(N)\in\{0,\varepsilon\}} e^{-\beta \varepsilon(N)} \right) = (1 + e^{-\beta \varepsilon})^N. \] (13.40)

Since this transformation is the key that makes the canonical formalism often easier than the microcanonical formalism, a more detailed explanation is in the following small lettered portion.

Suppose there are \( N \) lattice points. Each lattice point has several states \( a, b, c, \cdots \) with the corresponding ‘excitation energies’ \( \varepsilon(a), \varepsilon(b), \) etc. Since the total energy of the system, that is, the energy of the microstate, reads

\[ \mathcal{H} = \varepsilon(a_1) + \varepsilon(a_2) + \cdots + \varepsilon(a_N), \] (13.41)

the canonical partition function is computed as

\[ Z = \sum_{a_1,a_2,\cdots,a_N \in \{a,b,c,\cdots\}} e^{-\beta [\varepsilon(a_1) + \varepsilon(a_2) + \cdots]} \] (13.42)

Here, the summation is over all the possible combinations of the states of individual particles.

Since all the combinations appear once and only once, we can rewrite this as (see Fig. 13.2)

\[ Z = \left( \sum_{a_1 \in \{a,b,c,\cdots\}} e^{-\beta \varepsilon(a_1)} \right) \left( \sum_{a_2 \in \{a,b,c,\cdots\}} e^{-\beta \varepsilon(a_2)} \right) \cdots \left( \sum_{a_N \in \{a,b,c,\cdots\}} e^{-\beta \varepsilon(a_N)} \right) = Z_1^N, \] (13.43)

where

\[ Z_1 = \sum_{a_1 \in \{a,b,c,\cdots\}} e^{-\beta \varepsilon(a_1)} \] (13.44)
is the ‘canonical partition function’ of a single lattice point about its (internal) states. Notice that you cannot usually do this for the microcanonical approach, because not all the microstates appear in the computation of the microcanonical partition function $w$.

Thus, if ‘particles’ or ‘lattice points’ do not interact with each other, we can guess

$$
\sum_{\text{microstates}} = \left( \sum_{\text{one particle states}} \right)^N
$$

(13.45)

From this the Helmholtz free energy of the lattice with Schottky defects immediately follows:

$$
A = -N k_B T \log(1 + e^{-\beta \varepsilon}).
$$

(13.46)

We can get entropy by differentiation:

$$
S = -\frac{\partial A}{\partial T} = N k_B \log(1 + e^{-\beta \varepsilon}) + N \frac{\varepsilon}{T} \frac{e^{-\beta \varepsilon}}{(1 + e^{-\beta \varepsilon})}.
$$

(13.47)

Various other partition functions will be introduced in these lectures. As you will learn later, if you wish to study the thermodynamics of a system, any formalism will be OK, as long as the system is large enough (roughly speaking, if $\log N/N \ll 1$).

We have so far discussed the microcanonical and the canonical formalism. Let us check that the canonical result for $S$ agrees with the microcanonical result for this simple example. The microcanonical approach gives us

$$
S = -N k_B \left[ \frac{n}{N} \log \frac{n}{N} + \left( 1 - \frac{n}{N} \right) \log \left( 1 - \frac{n}{N} \right) \right],
$$

(13.48)

and

$$
\frac{n}{N} = \frac{1}{1 + e^{\beta \varepsilon}}.
$$

(13.49)
Combining both, we get

\[ S = -Nk_B \left[ \frac{1}{1 + e^{\beta \varepsilon}} \log \frac{1}{1 + e^{\beta \varepsilon}} + \frac{e^{\beta \varepsilon}}{1 + e^{\beta \varepsilon}} \log \frac{e^{\beta \varepsilon}}{1 + e^{\beta \varepsilon}} \right]. \]  

(13.50)

This indeed agrees with (13.47).

13.11 Another example: Dipoles on the honeycomb lattice with an easy direction
Consider a honeycomb lattice with \( N \) lattice points. At each lattice point is a dipole that can point in one of the three bond directions. If a dipole is along the easy direction, its energy is zero. If it points in other directions, its energy is \( \varepsilon (\geq 0) \). We assume dipoles do not interact.

![Image of honeycomb lattice with easy direction](image.png)

Figure 13.3: If a dipole is along the easy direction, it is energetically stabilized (by \( \varepsilon \)).

The canonical ensemble approach is easy. Using the idea explained around (13.40), we get

\[ Z(T) = (1 + 2e^{-\beta \varepsilon})^N, \]  

(13.51)

because at each lattice point one direction is an easy direction, and the other two have the energy penalty of \( \varepsilon \). The internal energy is

\[ E = -\frac{\partial \log Z}{\partial \beta} = N\varepsilon - \frac{2e^{-\beta \varepsilon}}{1 + 2e^{-\beta \varepsilon}}. \]  

(13.52)

With the aid of the principle of equal probability, we can ask the probability for a dipole to be in the leftward tilt:

\[ P(\text{left}) = \frac{e^{-\beta \varepsilon}}{1 + 2e^{-\beta \varepsilon}}. \]  

(13.53)

The probability to point in the easy direction is

\[ P(\text{easy}) = \frac{1}{1 + 2e^{-\beta \varepsilon}}. \]  

(13.54)
Dipole example with the microcanonical formalism

It is rather stupid to study this system with the microcanonical approach, but let us check that we can obtain the same result. Let \( A \) be the state of dipole tilting to the left, \( B \) to the right and \( C \) in the easy direction. Let \( n_X (X = A, B \text{ or } C) \) be the number of dipoles in state \( X \). \( N = n_A + n_B + n_C \). The number of microstates with definite \( n_A \) and \( n_B \) (\( n_C \) is determined) is (cf. the multinomial coefficients, see Appendix 2A)

\[
\frac{N!}{n_A!n_B!n_C!}.
\]  

(13.55)

To obtain the microstates with \( E = \varepsilon(N - n_C) = \varepsilon(n_A + n_B) \), we must collect all possible \( n_A \) and \( n_B \) compatible with the energy condition:

\[
w(\varepsilon(N - n_C)) = \sum_{n_A+n_B=N-n_C} \frac{N!}{n_A!n_B!n_C!} = \sum_{n_A+n_B=N-n_C} \frac{N!}{(N-n_C)!n_C!} \frac{(N-n_C)!}{n_A!n_B!}.
\]

\[
= \binom{N}{n_C} \sum_{n_A=0}^{N-n_C} \frac{(N-n_C)!}{n_A!n_B!} = \binom{N}{n_C} 2^{N-n_C}.
\]

You should have realized that this is easily obtained as follows. First we choose \( n_C \) sites to place easy-direction dipoles. There are \( \binom{N}{n_C} \) ways. Then, choose the remaining dipoles to tilt leftward or rightward. There are \( 2^{N-n_C} \) ways. Hence, we obtain the above result.

Thus, the entropy is

\[
S = k_B \log w(\varepsilon(N - n_C)) = -Nk_B \left[ \frac{n_C}{N} \log \frac{n_C}{N} + \left(1 - \frac{n_C}{N}\right) \log \left(1 - \frac{n_C}{N}\right)\right] + (N-n_C)k_B \log 2.
\]  

(13.56)

With the aid of the Gibbs relation (notice that \( dE = -\varepsilon dn_C \))

\[
\frac{1}{T} = \frac{\partial S}{\partial E} = -\frac{1}{\varepsilon} \frac{\partial S}{\partial n_C} = \frac{k_B}{\varepsilon} \log \frac{n_C}{N-n_C} + \frac{k_B}{\varepsilon} \log 2.
\]  

(13.57)

That is,

\[
e^{-\beta \varepsilon} = \frac{N-n_C}{2n_C}.
\]  

(13.58)

From this we obtain

\[
n_C = \frac{N}{1 + 2e^{-\beta \varepsilon}}.
\]  

(13.59)

The internal energy is

\[
E = \varepsilon(N - n_C) = N \varepsilon \frac{2e^{-\beta \varepsilon}}{1 + 2e^{-\beta \varepsilon}},
\]

(13.60)

which is identical to (13.52).
Appendix 13A. How to derive the principle of equal probability

You will see how important the law of large numbers is to establish statistical mechanics.

We first summarize fundamental properties of thermodynamic equilibrium states. We have already noted:

(O') If the equilibrium system is partitioned into two (approximately equal\textsuperscript{193}) 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 (Fig. 13.4).

We further know the fourth law or its direct consequence: Thermodynamic observables are obtainable from the partitioned system.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{thermodynamic_partition.png}
\caption{Thermodynamic quantities can be obtained from ‘pieces’ obtained by partitioning of an equilibrium state.}
\end{figure}

Thus, we may conclude that thermodynamic equilibrium states are partitioning-rejoining invariant.

Although usually not stated clearly, we know one more fact:

(Y)\textit{ Invariance under thermal contact} of equilibrium states: Any equilibrium state of a thermally isolated system has a heat bath (\textit{intrinsic heat bath}) such that thermal contact with it does not alter its thermodynamic state (Fig. 13.5).

(O’) and (Y) imply that the following procedure keeps thermal equilibrium intact (Fig. 13.6).

(i) Partition a thermally isolated equilibrium system into (macroscopic) pieces.

(ii) Attach each piece to its private intrinsic heat bath for a while, and then again thermally isolate it.

(iii) Re-join all the pieces as before to reconstruct the whole piece.

\textsuperscript{193}This requirement is only to avoid extreme cases in which one part contains not macroscopic number of particles.
Thermally isolated system
= 
Intrinsic heat bath

Figure 13.5: There is a heat bath that does not destroy a given equilibrium system upon thermal contact.

van Hove partition

Contact with individual heat baths

rejoining

Figure 13.6: An equilibrium system may be replaced by statistically independent pieces to obtain thermodynamic quantities.

The above procedure does not alter thermodynamic observables of the system. Thus, a macroscopic system in thermal equilibrium may be replaced with a collection of many statistically independent mechanical systems, if we are interested in thermodynamic quantities. In other words, thermodynamics observables are such physical quantities that are quite insensitive to subtle correlations (as quantum mechanics implies) among portions of a system. Therefore, we may use a brutal procedure to compute them.

How many such independent pieces can we find in an ordinary macroobject? 1 mm$^3$ is big enough from the molecules’ point of view, so in a cube with 10 cm edge, we can easily expect more than $10^6$ macroscopic subsystems; we may safely use the law of large numbers, so thermodynamic quantities obtained from the partitioned system and the actual values are quite close.

The reader may even expect that to compute thermodynamic quantities, we may use the principle of equal probability = all the microstates compatible with a given thermodynamic state are equally probably sampled.

The logic to demonstrate this statement is as follows:
(i) The direct product model may be used as a micro-model of a thermodynamic system.
(ii) Then, the asymptotic equipartition (see below) implies that the correct expectation values are obtained, even if we assume all the energy states are equally probably
(iii) Thus, to obtain thermodynamics, we may assume that all the compatible microstates are equally probable. This is called the principle of equal probability traditionally assumed to obtain statistical mechanics.

The asymptotic equipartition is nothing but the law of large numbers.

\[
\frac{1}{n} \log P(x_1, \cdots, x_n) = \frac{1}{n} \sum_i \log P(x_i) \converges to -s
\]  

(13.61)

in the large \( n \) limit. Here \( s \) is the entropy per piece. That is, the asymptotic equipartition law:

\[
P(x_1, \cdots, x_n) = e^{\frac{-ns \pm o[n]}{n}}
\]  

(13.62)

holds independent of the actual microstate \( \{x_1, \cdots, x_n\} \).
Q13-1. There is a 2D square lattice with $M$ lattice points. On each lattice point is a magnetic moment that can point only in the lattice bond directions (4 directions as illustrated), but the $\pm y$ directions are the easy directions: if the dipole is along the $y$-axis, it is stable, that is, the energy of the dipole along the $x$-axis is $\varepsilon (> 0$ more energy) and that along the $y$-axis is zero. We do not pay attention to the kinetic energy of the system. You may ignore the interactions among dipoles.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure13.7.png}
\caption{Each dipole can point only 4 direction along the lattice bonds.}
\end{figure}

(1) What is the canonical partition function of the system (the temperature is $T$)?

(2) What is the average energy per dipole?

(3) Compute the entropy $S(T)$ per dipole. What is the difference $S(\infty) - S(0)$? How many bits is this? Is this consistent with the intuitive interpretation of entropy per molecule as the number of YES-NO questions?

(4) Compute the ‘microcanonical partition function’ $w(N\varepsilon) (0 \leq N \leq M)$.

(5) Show that the entropy you computed from the microcanonical scheme (Boltzmann’s principle) and the result (3) agree. [Compute $N/M$ as a function of $T$ (use $1/T = \partial S/\partial E$) and get rid of $N/M$ from the formula to obtain the result of (3).]

Soln. (1) Needless to say, you can start from the very definition of the canonical partition function, BUT notice that if you collect all the microstates (= microscopic states = mechanically describable whole-system states), all the states of microscopic entities (molecules, etc., in our case dipoles sitting on the lattice points) appear once and only once (recall Fig. 13.2). Therefore, to construct the partition function for the whole lattice, we study all the states of each microscopic entity to make their individual canonical partition functions (in our case $1 + 1 + e^{-\beta\varepsilon} + e^{-\beta\varepsilon}$) and multiply them over the whole lattice:

$$Z(T) = (2 + 2e^{-\beta\varepsilon})^M.$$  

(2) Using the Gibbs relation, we get

$$\frac{E}{M} = -\frac{1}{M} \frac{\partial \log Z}{\partial \beta} = \frac{\varepsilon e^{-\beta\varepsilon}}{1 + e^{-\beta\varepsilon}}.$$
(3) Since $A = E - TS$, $S = (E - A)/T$

$$\frac{S}{M} = k_B \log(2 + 2e^{-\beta\varepsilon}) + \frac{\varepsilon}{T} \frac{e^{-\beta\varepsilon}}{1 + e^{-\beta\varepsilon}}.$$ 

$S(0) = k_B \log 2$ (notice that for any $n$, $x^n e^{-x} \to 0$ in the $x \to \infty$ limit) and $S(\infty) = k_B \log 4$, so $\Delta S$ is just 1 bit. At $T = 0$ the dipoles are always along $y$ (2 directions), but at $T = \infty$ they can evenly assume 4 directions (i.e., along $x$ and $y$ directions). Thus, if we ask one yes-no question (“Is it along $y$?”) we can reduce the uncertainty in the equilibrium state at $T = \infty$ to that at $T = 0$. In other words, to identify the state of a dipole in the $T \to \infty$ limit we need 2 bits (2 questions), because we must find 1 particular state out of 4 possibilities. In the $T \to 0$ limit all the dipoles are along the easy direction, so there are only 2 choices for each dipole. Therefore, we need only one question to pinpoint the state of a dipole. $\Delta S$ just corresponds to the difference in the numbers of questions we must ask.

(4) You can immediately obtain

$$w(N\varepsilon) = \binom{M}{N} 2^M,$$

because 2 choices along $x$ and along $y$ can be selected without affecting the system energy.

A more pedestrian way (which I do not recommend) is to introduce $n_1, n_2, n_3$ and $n_4$ pointing respectively $+x$, $-x$, $+y$ and $-y$. $N = n_1 + n_2$ ($x$-direction) and $M - N = n_3 + n_4$ ($y$-direction):

$$w(N\varepsilon) = \sum_{n_1=0}^{N} \sum_{n_3=0}^{M-N} \frac{M!}{n_1!(N-n_1)!n_3!(M-N-n_3)!},$$

but an easy reorganization is: (i) choose $N$ parallel $x$ dipoles, and then (ii) count the number of ways to point $+$ and $-$ directions:

$$w(N\varepsilon) = \binom{M}{N} \sum_{n_1=0}^{N} \sum_{n_3=0}^{M-N} \binom{N}{n_1} \binom{M-N}{n_3} = \binom{M}{N} 2^{N_M-N}.$$

This is just the answer above.

(5) Thanks to Boltzmann

$$\frac{S}{M} = -k_B \left[ \frac{N}{M} \log \frac{N}{M} + \left( 1 - \frac{N}{M} \right) \log \left( 1 - \frac{N}{M} \right) \right] + k_B \log 2.$$
I strongly urge you to learn the following by heart:

\[
\log \left( \frac{M}{N} \right) = -M \left[ \frac{N}{M} \log \frac{N}{M} + \left(1 - \frac{N}{M} \right) \log \left(1 - \frac{N}{M} \right) \right].
\]

Since

\[
\frac{1}{T} = \frac{k_B}{\varepsilon} \frac{\partial S}{\partial N} = -\frac{k_B}{\varepsilon} \log \frac{N}{M - N},
\]

we have

\[
N = \frac{M}{1 + e^{-\beta \varepsilon}}.
\]

Computing \(N/M\) and using it in the entropy formula above, we get

\[
\frac{S}{M} = -k_B \left[ \frac{N}{M} \log \frac{1}{1 + e^{-\beta \varepsilon}} + \left(1 - \frac{N}{M} \right) \log \left(\frac{e^{-\beta \varepsilon}}{1 + e^{-\beta \varepsilon}}\right) \right]
\]

\[
= k_B \left[ \frac{N}{M} \log(1 + e^{-\beta \varepsilon}) + \left(1 - \frac{N}{M} \right) \log(1 + e^{-\beta \varepsilon}) \right] + 2k_B \log 2 - k_B \frac{e^{-\beta \varepsilon}}{1 + e^{-\beta \varepsilon}} \log \left(\frac{1}{e^{-\beta \varepsilon}}\right)
\]

\[
= k_B \log(2 + 2e^{-\beta \varepsilon}) + \frac{\varepsilon}{T} \frac{e^{-\beta \varepsilon}}{1 + e^{-\beta \varepsilon}}.
\]

**Q13-2.** We wish to obtain the chemical potential of (classical) ideal gas. Since we know \(G = N\mu\), we wish to obtain the Gibbs free energy directly statistical-mechanically. We know

\[
-G \frac{T}{T} = -\frac{A + PV}{T} = S - \frac{E}{T} - \frac{PV}{T}.
\]

(13.63)

Therefore, we easily guess

\[
-G \frac{T}{T} = k_B \log Y(T, P),
\]

(13.64)

where

\[
Y(T, P) = \int dV Z(T, V, X) e^{-\beta PV} = \int dE \int dV w(E, V, CX) e^{-(E + PV)/k_B T}
\]

\[
= \int dE \int dV e^{[S(E) - E/T - PV/T]/k_B}.
\]

(13.65)

(1) Compute \(Y\) for a classical ideal gas of mass \(m\). You may use the already obtained canonical partition function \(Z\) for the same system.

(2) Show indeed \(G\) obtained from \(Y\) (that is, \(G\) directly obtained by statistical mechanics) agrees with \(A + PV\) computed by the canonical formalism (from \(Z\); you
can use already given results) + thermodynamics (the definition of $G$).

**Soln.** (1)

$$Z = \frac{1}{N!} \left[ \left( \frac{2\pi mk_B T}{\hbar^2} \right)^{3/2} \right]^N V^N,$$

so we need

$$\int dV \frac{1}{N!} V^N e^{-\beta PV} = \frac{1}{N!(\beta P)^{N+1}} \int dx x^N e^{-x} = \frac{1}{N!(\beta P)^{N+1}} \Gamma(N + 1) = \left( \frac{k_B T}{P} \right)^N.$$

Therefore, (I ignore the difference between $N$ and $N + 1$)

$$Y(T, P) = \left[ \left( \frac{2\pi m}{\hbar^2} \right)^{3/2} \left( \frac{k_B T}{P} \right)^{5/2} \right]^N.$$

(2) Comparing $Y$ and $Z$ we obtain

$$Y = \left( \frac{N!}{V^N} \right) Z \left( \frac{k_B T}{P} \right)^N = Z \left( \frac{N}{eV} \right)^N \left( \frac{k_B T}{P} \right)^N = Z \left( \frac{N k_B T}{PV} \right)^N e^{-N}.$$

Therefore,

$$G = -k_B T \log Y = -k_B T \log Z - N k_B T \log \left( \frac{N k_B T}{PV} \right) + N k_B T.$$

Our system is an ideal gas, so $N k_B T / PV = 1$. Therefore, the above formula reads

$$G = A + N k_B T = A + PV.$$

Here, we relied on a particular model to confirm this relation, but we can prove this very generally roughly as follows (you will see in the next Lecture that our logic is always the same and the simple one)

$$\max_V [Z(T, V, X)e^{-\beta PV}] \leq Y = \int dV Z(T, V, X)e^{-\beta PV} \leq cV \max_V [Z(T, V, X)e^{-\beta PV}],$$

where $c$ is a numerical factor of order 1 (you can choose it as $N$ to be (very) safe).

**Q13-3.** It is convenient to study a magnet under constant magnetic field $H$ rather
than under constant magnetization $M$. Then, the thermodynamic independent variables should be $T, H$ rather than $E, M$. Therefore, we should use the ‘magnetic-counterpart’ of the Gibbs free energy $\tilde{G} = A - HM$ rather than the Helmholtz free energy $A$ itself.\(^{194}\) Notice that $d\tilde{G} = -sdT - MdH$. 

(1) Let $H_0$ be the Hamiltonian (energy) of the magnetic system. Define an appropriate partition function $Q$ that directly gives

$$\tilde{G} = -k_B T \log Q(T, H). \quad (13.66)$$

You must define $Q$ in terms of $Z(T, M)$. [Needless to say, 12-2 is a hint, but be careful about the sign.]

(2) Consider a collection of $N$ magnetic dipoles $\mu_i$ whose magnetization can be written as $M = \sum \mu_i$. The magnetic moments can point only up or down state (of value $\pm \mu$) aligned to the magnetic field direction (say, along the z-axis; its potential energy is $\mp H\mu$). If we ignore the kinetic energy of the magnetic moments, the system Hamiltonian of the noninteracting magnetic moments is just 0 (no energy associated).\(^{195}\) Compute $Q$ and obtain the magnetization as a function of $T$ and $H$. (3) Show that $M$ obtained from the microcanonical approach agrees with the result of (2).

**Soln.** (1)

$$-\frac{\tilde{G}}{T} = -\frac{A}{T} - \frac{HM}{T} = k_B \log \sum_M Z(T, M) e^{\beta HM}$$

That is,

$$\tilde{G} = -k_B T \log Q$$

with

$$Q = \sum_M e^{-\beta(H_0 - HM)}.$$

However, in our case the system energy is zero ($H_0 = 0$), so $Z(T, M) = w(0, M)$ or

$$Q = \sum e^{\beta HM} = \sum_{\mu_i \in \{\mu, -\mu\}} e^{\beta H \sum \mu_i}.$$
where the summation is over all the microstates.

(2) \[ Q = \sum_{M} e^{\beta H M} = \left(2 \cosh \frac{H \mu}{k_B T}\right)^N. \]

Hence, \[ \tilde{G} = -Nk_B T \log \left(2 \cosh \frac{H \mu}{k_B T}\right). \]

From this \[ M = -\frac{\partial \tilde{G}}{\partial H} \bigg|_{T} = N\mu \tanh \frac{H \mu}{k_B T}. \]

(3) The microcanonical partition function we need is \( w(0, M) \). Let \( N_+ \) be the number of up-spins. Then,

\[ w = \binom{N}{N_+}, \]

so \[ S = -k_B N \left[ \frac{N_+}{N} \log \frac{N_+}{N} + \left(1 - \frac{N_+}{N}\right) \log \left(1 - \frac{N_+}{N}\right) \right]. \]

The number of down-spins is \( N - N_+ \), so \( M = \mu (2N_+ - N) \). Therefore, \( N_+ = \frac{(M + \mu N)}{2\mu} \). From \( dS = -(H/T) dM \) for our system (recall we cannot change \( E \equiv 0 \)),

\[ \frac{H}{T} = -\frac{\partial S}{\partial M} \bigg|_E = -\frac{1}{2\mu} \frac{\partial S}{\partial N_+} \bigg|_E = \frac{k_B}{2\mu} \log \frac{N_+/N}{1 - (N_+/N)}, \]

that is,

\[ N_+ = \frac{e^{\beta \mu H}}{1 + e^{2\beta \mu H}} = \frac{N}{1 + e^{-2\beta \mu H}}. \]

From this, we get the same result:

\[ M = N\mu \frac{1 - e^{2\beta \mu H}}{1 + e^{2\beta \mu H}} = N\mu \tanh \frac{\mu H}{k_B T}. \]

Which do you think is easier?
14 Classical ideal gas and quantum-classical correspondence

Summary
* All ensembles are equivalent. You can use any that you can compute most easily.
* You must recognize that the ensemble equivalence is, in its essence, a crude relation.
* We compute the classical ideal gas partition function. We must treat all gas particles indistinguishable.
* The canonical partition function of a classical fluid system reads

\[ Z = \frac{1}{N!\hbar^{3N}} \int d\Gamma e^{-\beta H}. \]

* The classical-quantum partition function relation may be understood as the requirement that the partition function must be dimensionless.
* Equipartition of energy and classical specific heat is studied with the aid of the canonical formalism. Quantization turns out to be mandatory.

Key words
ensembles, ensemble equivalence, Gibbs paradox, indistinguishability, Frenkel defect, equipartition of energy

What you should be able to do
* Get the canonical partition function of a classical gas with dimensional analysis.
* Be able to explain what the ensemble equivalence means, and the condition that we can use any ensemble.
* Be able to use the equipartition formula to compute simple averages.
* Be able to explain why the specific heat of a diatomic gas is not as big as expected classically.

14.1 Review: ensemble equivalence
We have learned two formalisms to do equilibrium statistical mechanics:

\[ S = k_B \log w(E, X) \quad \text{microcanonical formalism}, \quad (14.1) \]
\[ A = -k_B T \log Z(T, X) \quad \text{canonical formalism}. \quad (14.2) \]

These formalisms are equivalent if \( \log N/N \ll 1 \). The meaning of ‘equivalence’ is: The free energy computed from \( S \) according to (14.1) agrees with that computed
directly from statistical mechanics according to (14.2); The entropy computed from $A$ according to (14.2) agrees with $S$ directly computed statistical mechanically from (14.1). In short, you may use any ‘ensemble.’ Here, ‘ensemble’ means the collection of microscopic states with a definite summation rule (or probability assignment, if we use the principle of equal probability).

14.2 Derivation of Maxwell’s distribution

The principle of equal probability implies that the probability for a particular particle, called 0 here, to have energy $\varepsilon$ in a macrosystem is given by

$$P(\varepsilon) = \frac{1}{Z} \sum_{\gamma \in \Gamma_{\varepsilon}} e^{-\beta H(\gamma)}, \quad (14.3)$$

where $\Gamma_{\varepsilon}$ is the totality of microstates compatible with particle 0 to have energy around $\varepsilon$. The summation is an informal expression of summing or integration over all the appropriate microstates.\(^{196}\)

For an ideal gas, $H(\gamma)$ consists of a sum of the form $\sum_i \varepsilon_i$, where $\varepsilon_i$ is the energy for $i$-th particle. The variables are statistically independent; this is the meaning of the ideality. Therefore, obviously,

$$P(\varepsilon) \propto e^{-\beta \varepsilon}. \quad (14.4)$$

Thus, we recover Maxwell’s distribution function, because $\varepsilon = m v^2 / 2$:

$$P(v) \propto e^{-m v^2 / 2k_B T}. \quad (14.5)$$

14.3 Frenkel defect

Let us study an example: the Frenkel defects. As illustrated in Fig. 14.1 particles leave their original lattice points and wander into non-lattice positions (interstitial positions). There are $N$ lattice points and $M$ interstitial points. There are total $N$ particles and $n$ particles leave their lattice points and move into interstitial sites. There is an energy cost of $\varepsilon$ to leave a lattice point to move to an interstitial site.

\(^{196}\)Clearly notice that we are not dealing with individual microstates of a macrosystem, because $\Gamma_{\varepsilon}$ contains macroscopically many microstates.
The system energy is $E = n\varepsilon$. Thus, (we write $w(E)$ as $w(n)$)

$$w(n) = \binom{N}{n} \binom{M}{n},$$

(14.6)

and the canonical partition function reads

$$Z(T) = \sum_{n=0}^{N} w(n) e^{-\beta n\varepsilon} = \sum_{n=0}^{N} \binom{N}{n} \binom{M}{n} e^{-\beta n\varepsilon}.$$  

(14.7)

Unfortunately, this cannot be summed in a closed form. However, in this case, it is very easy to prove the ensemble equivalence as follows.

### 14.4 Equivalence of microcanonical and canonical ensembles

Since all the summands are positive, the following inequalities are obvious:

$$\max_n \left[ w(n) e^{-\beta n\varepsilon} \right] \leq Z(T) \leq N \max_n \left[ w(n) e^{-\beta n\varepsilon} \right].$$  

(14.8)

On the other hand, we have

$$\max_n \left[ w(n) e^{-\beta n\varepsilon} \right] = \exp \left[ \frac{1}{k_B} \max_E (S - E/T) \right],$$

(14.9)

but the ‘honest’ definition of the Helmholtz free energy is

$$-A = \max_E [TS - E].$$

(14.10)

Therefore, (14.8) reads

$$-\frac{A}{k_B T} \leq \log Z(T) \leq -\frac{A}{k_B T} + \log N.$$  

(14.11)
A is an extensive quantity, so it is of order $N$. Therefore, if you can ignore $\log N/N$ as very small, $-k_B T \log Z(T)$ (the free energy directly obtained by the canonical formalism of statistical mechanics) and the free energy obtained (using thermodynamics, i.e., (14.10)) from entropy (which is computed statistical-mechanically with the aid of the microcanonical approach) are indistinguishable.

Although our demonstration of the ensemble equivalence, in the present case the equivalence of the canonical and microcanonical formalisms, relies on a particular example, the logic we have employed is identical to the key logic to demonstrate the ensemble equivalence generally and rigorously: $Z$ is sandwiched between the maximum term and the maximum term $\times$ something proportional to $N$.

You must clearly recognize that the estimations used in statistical mechanics are very crude. That is why the results are general and very stable. Also you must clearly recognize that to compute $Z$ is equivalent to estimating its maximum summand.

14.5 Frenkel defect, microcanonical approach
Let us continue the Frenkel defect problem. Let us compute entropy using Boltzmann’s principle.

$$S = k_B \log \binom{N}{n} \binom{M}{n}. \quad (14.12)$$

We use Stirling’s approximation, or

$$\log \left( \frac{A}{B} \right) = -A \left[ \frac{B}{A} \log \frac{B}{A} + \left( 1 - \frac{B}{A} \right) \log \left( 1 - \frac{B}{A} \right) \right]. \quad (14.13)$$

We obtain

$$S/k_B = -N \left[ \frac{n}{N} \log \frac{n}{N} + \left( 1 - \frac{n}{N} \right) \log \left( 1 - \frac{n}{N} \right) \right] - M \left[ \frac{n}{M} \log \frac{n}{M} + \left( 1 - \frac{n}{M} \right) \log \left( 1 - \frac{n}{M} \right) \right]. \quad (14.14)$$

We need temperature, so we use the Gibbs relation $dS = (1/T)dE + (P/T)dV + \cdots$:

$$\frac{1}{T} = \frac{\partial S}{\partial E} \bigg|_V, \quad (14.15)$$

but $dE = \varepsilon dn$,\(^{197}\) so

$$\frac{\varepsilon}{k_B T} = \frac{1}{k_B} \frac{dS}{dn} = \log \frac{1 - \frac{n}{N}}{\frac{n}{N}} + \log \frac{1 - \frac{n}{M}}{\frac{n}{M}} = \log \frac{(N - n)(M - n)}{n^2}. \quad (14.16)$$

\(^{197}\) $S$ is now a function of $E$, but we do not write it explicitly.
Here, notice that when you differentiate \( \log \left( \frac{N}{n} \right) \) wrt \( n \), virtually you have only to differentiate the factors outside \( \log \). Usually \( n \) is small, so we obtain

\[
e^{-\beta \varepsilon} = \frac{n^2}{NM}.
\]  

(14.17)

From this we can write \( S \) in terms of \( T \).

### 14.6 Classical ideal gas and de Broglie wavelength

The classical ideal gas is characterized by the total absence of quantum effect: here, quantum effect means that the particles can delocalize. Consider a gas consisting of \( N \) identical noninteracting particles. To ignore all quantum effects, the average de Broglie wave length of each particle must be much smaller than the average interparticle distance. The de Broglie wave length \( \lambda \) may be estimated as

\[
\lambda \sim \frac{\hbar}{\sqrt{mk_B T}},
\]

(14.18)

where \( m \) is the mass of the particle, and \( \hbar \) is Planck’s constant. This estimate is due to \( \lambda = h/p = \hbar / \sqrt{2mK} \) and \( K \sim k_B T \), where \( p \) is the representative value of the magnitude of the momentum of a particle, and \( K \) is the representative value of the one-particle kinetic energy. The mean particle distance is \( \sqrt[3]{V/N} \), so the condition we want is \( \sqrt[3]{V/N} \gg \lambda \), or

\[
\frac{N}{V} \ll \left( \frac{mk_B T}{\hbar^2} \right)^{3/2}.
\]

(14.19)

When this inequality is satisfied, we say the gas is classical.\(^{198}\)

Since there are no interactions among particles, each particle cannot sense the density. Consequently, the internal energy of the system must be a function of \( T \) only: \( E = E(T) \). This is a good characterization of ideal gases.

### 14.7 Single particle states

Let us first compute the number of microscopic states allowed to a single particle in a box of volume \( V \). To this end we solve the Schrödinger equation in a cube with edges of length \( L \):

\[
-\frac{\hbar^2}{2m} \Delta \psi = E\psi;
\]

(14.20)

\(^{198}\)Notice that the dynamics of internal degrees of freedom such as vibration and rotation need not be classical as we will see in Section 23.
\( \Delta \) is the Laplacian, and a homogeneous Dirichlet boundary condition \( \psi = 0 \) at the wall is imposed. As is well-known, the eigenfunctions are:

\[
\psi_k \propto \sin k_x x \sin k_y y \sin k_z z
\]  

(14.21)

with the following quantization condition due to the boundary condition:

\[
k \equiv (k_x, k_y, k_z) = \frac{\pi}{L} (n_x, n_y, n_z) \equiv \frac{\pi}{L} n.
\]  

(14.22)

Here, \( n_x, \cdots \) are positive integers, 1, 2, \cdots. The eigenfunction \( \psi_k \) belongs to the eigenvalue (energy) \( \hbar^2 k^2 / 2m \).

The number of states with wave number vectors \( k \) in the range \( k \) to \( k + dk \) is

\[
\#\{k \mid k < |k| < k + dk\} = \# \left\{ n \mid \frac{L}{\pi} k < |n| < \frac{L}{\pi} (k + dk) \right\}
\]

\[
= \left( \frac{1}{8} 4\pi n^2 dn = \right) \frac{1}{8} \frac{L^3}{\pi^3} 4\pi k^2 dk = \frac{1}{2\pi^2} V k^2 dk.
\]  

(14.23)

The factor 1/8 is required because the relevant \( k \) are only in the first octant (all the components must be positive).

### 14.8 Classical ideal gas: single particle canonical partition function

Now we can compute the canonical partition function for a single particle using its definition:

\[
Z_1 = \sum_{n_x>0, n_y>0, n_z>0} \exp(-\beta E)
\]

(14.24)

\[
= \int_k \# \left\{ n \mid \frac{L}{\pi} k < |n| < \frac{L}{\pi} (k + dk) \right\} \exp(-\beta k^2 \hbar^2 / 2m)
\]

(14.25)

\[
\approx \frac{1}{8} \frac{V}{\pi^3} \int_0^{\infty} 4\pi k^2 dk \exp(-\beta k^2 \hbar^2 / 2m). \quad \text{(14.26)}
\]

The integration is readily performed. (14.26) is

\[
Z_1(V) = V \frac{1}{8\pi^3} \int_{-\infty}^{\infty} dk_x e^{-k_x^2 \hbar^2 / 2mk_BT} \int_{-\infty}^{\infty} dk_y e^{-k_y^2 \hbar^2 / 2mk_BT} \int_{-\infty}^{\infty} dk_z e^{-k_z^2 \hbar^2 / 2mk_BT}, \quad \text{(14.27)}
\]

and we know

\[
\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\frac{\pi}{a}}. \quad \text{(14.28)}
\]
Therefore,
\[
Z_1(V) = V \frac{1}{8\pi^3} \left( \frac{2\pi mk_BT}{\hbar^2} \right)^{3/2} = V \left( \frac{1}{4\pi^2} \right)^{3/2} \left( \frac{8\pi^3 mk_BT}{\hbar^2} \right)^{3/2}.
\] (14.29)

That is,
\[
Z_1(V) = V \left( \frac{2\pi mk_BT}{\hbar^2} \right)^{3/2}.
\] (14.30)

The important point of this result is that \( Z_1 \propto V \).

14.9 Classical ideal gas: Gibbs paradox

According to our preliminary discussion around (13.45) the partition function \( Z \) of the whole ideal gas system consisting of \( N \) identical particles should read
\[
"Z = Z_1^N".
\] (14.31)

This implies
\[
A(N,V) = -Nk_BT \log Z_1(V).
\] (14.32)

Now, prepare two identical systems each of volume \( V \) with \( N \) particles. The free energy of each system is given by \( A(N,V) \). Next, combine these two systems to make a single system. The resultant system has \( 2N \) particles and volume \( 2V \), so its free energy should be \( A(2N,2V) \). The fourth law of thermodynamics requires that
\[
A(2N,2V) = 2A(N,V).
\] (14.33)

Unfortunately, as you can easily check, this is not satisfied by (14.32). \( Z_1 \propto V \) is the key feature, so let us write \( Z_1 = cV \) with a positive constant \( c \). Then, \( Z = (cV)^N \), so indeed
\[
\log(c2V)^{2N} = 2\log(cV)^N + \log 2^{2N} \neq 2\log(cV)^N.
\] (14.34)

Thus we must conclude (14.31) is wrong. This is the famous Gibbs paradox.

Since the fourth law is an empirical fact, we must correct (14.31) as
\[
Z = f(N)Z_1^N = f(N)(cV)^N,
\] (14.35)

where \( f(N) \) is as yet an unspecified function of \( N \). The fourth law demands (14.33):
\[
\log f(2N) + 2N \log(c2V) = 2\log f(N) + 2N \log(cV).
\] (14.36)

That is,
\[
\log f(2N) + 2N \log 2 = 2\log f(N)
\] (14.37)
or
\[ f(N)^2 = 2^N f(2N) \] (more generally, \( f(N)^\alpha = \alpha^{\alpha N} f(\alpha N) \)). \hspace{1cm} (14.38)

The general solution to this functional equation is (set \( \alpha = 1/N \); recall Stirling’s formula \( (N/e)^N \approx N! \))
\[ f(N) = \left( \frac{f(1)}{N} \right)^N \propto (N!)^{-1}. \] \hspace{1cm} (14.39)

Thus, thermodynamics forces us to write
\[ Z = \frac{1}{N!} Z_i^N, \] \hspace{1cm} (14.40)

where we have discarded the unimportant multiplicative factor.

Therefore, the canonical partition function for a classical ideal gas reads
\[ Z_{\text{ideal}} = \frac{V^N}{N!} \left( \frac{2\pi m k_B T}{\hbar^2} \right)^{3N/2} = \left[ \frac{Ve}{N} \left( \frac{2\pi m k_B T}{\hbar^2} \right)^{3/2} \right]^N. \] \hspace{1cm} (14.41)

14.10 Why does a gas require \( 1/N! \), but does a lattice system not?
For a lattice system we considered in the preceding lecture, a spatial pattern consisting of states at individual lattice points is identified as a microstate (see Fig. 14.2 (1) and (2)). In this case, the particles sitting at the lattice points are identical chemical species (atom or molecule) and we never pay any attention to the arrangement of particles on the lattice. If these particles are marbles, then their arrangements on the lattice can distinguish microstates, but since we do not do that, we have already assumed that all the particles are (combinatorially) indistinguishable.

Even for a gas system a microstate corresponds to a pattern of particle positions and momentum vectors (see Fig. 14.2 (3) or (4)), since as we will learn later, quantum-mechanically identical particles are indistinguishable combinatorially in contrast to marbles. However, if you use classical mechanics to describe a gas we must name particles to describe them separately. Consequently, the identical patterns with differently named particles look as if they are distinct microstates. In Fig. 14.2 (5) and (6) are with an identical pattern (3) so they must represent an identical microstate, but due to different namings they are handled as distinct microstates. If there are \( N \) gas particles, the number of microstates are multiplied with \( N! \) due to namings of particles, so to correct this overcounting \( 1/N! \) must be multiplied.
harmonic oscillator lattice
numbers are quantum numbers itinerating gas particles

Figure 14.2: Left: A lattice system we studied in the preceding lecture; Right: a particle system. For the lattice system, a microstate is identified with a pattern on the whole lattice that is an arrangement of states at all lattice points. Thus, (1) and (2) are distinct microstates since they have distinct spatial patterns. Notice that the particles sitting at lattice points are already indistinguishable. The situation does not vary very much for gasses. A pattern consisting of positions and momentum vectors of the particles corresponds to a single microstate. Thus, (3) and (4) are distinct microstates. Here, what particle is assigned to what position does not matter. In contrast to the lattice system, however, when the partition function is written down, particles have definite names as a, ⋅⋅⋅, i and are distinguishable. Consequently, a certain single microstate of \( N (= 9 \text{ in the figure}) \) particles are distinguished into \( N! \) apparently distinct microstates due only to the naming of the particles. For example, due to distinct names of the particle variables required in the classical description, the microstate (3) may be written as (5), (6) or other \( N! \) distinct configurations.

14.11 Classical canonical partition function of particle system
Let’s recap. If we ‘honestly count’ the number of quantum states to obtain the microcanonical partition function (i.e., \( w(E, X) \)), we see (→14.30)

\[
Z_1 = \frac{1}{\hbar^3} \int d\mathbf{r} \int p e^{-p^2 / 2mk_BT}.
\]  

(14.42)

Therefore, the canonical partition function in terms of the phase integral reads

\[
Z = \frac{1}{\hbar^{3N} N!} \int d\Gamma e^{-\beta \sum \mathbf{p}_i^2 / 2m},
\]  

(14.43)

where \( d\Gamma = d\mathbf{q}_1 d\mathbf{q}_2 \cdots d\mathbf{q}_N d\mathbf{p}_1 d\mathbf{p}_2 \cdots d\mathbf{p}_N \) is the phase volume element.

The prefactors in front of the phase integral are determined while considering a particular system (the classical ideal gas), so you may think they are rather ad hoc.
However, the factor $1/N!$ comes from the indistinguishability of the particles, so as long as particles are distinguished in the description of the system, this should not be peculiar to ideal gases. How about $h^{3N}$?

$Z$ appears in log, so it must be dimensionless (if not, the free energy shifts according to the choice of units, for example). Therefore, in front of the integral whose dimension is $(\text{action})^{3N}$ ($[pq] = M(L/T)L = M(L/T)^2 \times T^{199}$), we must have a factor killing this dimension. The most fundamental quantity in physics that has the dimension of action is $h$. Since we do not expect that the factor is idiosyncratic to the ideal gas, it is natural to expect $1/h^{3N}$ to appear. Therefore, we define the canonical partition function in terms of the phase integral as follows:

$$Z = \frac{1}{h^{3N}N!} \int d\Gamma e^{-\beta H}. \quad (14.44)$$

This relation can be rigorously demonstrated by semi-classical analysis.

14.12 Dimensional analysis of ideal gases

We did a lot of calculation to get $Z_1$ quantum mechanically. However, dimensional analysis almost gives you the same result, or, actually, since we may ignore any constant factor, we get the correct result by dimensional analysis alone. Our starting point is $Z_1 \propto V$. This is very natural. To make a dimensionless quantity, we need another quantity with the dimension of volume, or rather, if we can find a quantity with the dimension of length, we can use it to make $Z$ dimensionless. What length scales do we have in this problem? The system size (the box size $L$) is certainly relevant, but we have already used it $V = L^3$. There is one more length scale, which we have already discussed: the de Broglie wave length $\lambda \sim \sqrt{mk_BT/h^2}$. Therefore,

$$Z_1 \propto \frac{V}{\lambda^3} = V \left(\frac{mk_BT}{h^2}\right)^{3/2}. \quad (14.45)$$

Compare this with (14.30). Needless to say, $1/h^3$ appears naturally.

14.13 Generalization of equipartition of kinetic energy

We already know the equipartition of kinetic energy for an ideal gas with the aid of the kinetic theory of gases, e.g.,

$$\left\langle \frac{1}{2}m\dot{x}_x^2 \right\rangle = \frac{1}{2}k_BT. \quad (14.46)$$

\footnote{Action is energy times time. Recall $h\nu$ is energy.}
Let us demonstrate, with the aid of the canonical formalism, a general theorem that implies the above formula and that is applicable to any classical systems.

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

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = k_B T \delta_{ij}, \quad (14.47)$$

where the average is over the canonical distribution. Indeed,

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = \frac{1}{Z} \int d\Gamma x_i \left[ -k_B T \frac{\partial}{\partial x_j} e^{-\beta H} \right], \quad (14.48)$$

$$= -\frac{1}{Z} k_B T x_i e^{-\beta H} \bigg|_{|x| \to \infty} + \frac{1}{Z} k_B T \int d\Gamma \frac{\partial x_i}{\partial x_j} e^{-\beta H}. \quad (14.49)$$

Here, the first term due to an integration by parts must vanish, so $H$ must increase sufficiently fast in the large variable limit. For example, if a system is spatially confined (by a potential well), certainly this is true for the spatial coordinates.

14.14 Equipartition of kinetic energy again

From (14.47) we obtain the law of equipartition of energy for classical kinetic energy such as (no summation convention implied)

$$\langle \frac{p_i^2}{2m} \rangle = \frac{1}{2} k_B T, \quad (14.50)$$

or

$$\langle \frac{L_i^2}{2I_i} \rangle = \frac{1}{2} k_B T, \quad (14.51)$$

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 spatial position of a particle is governed by a harmonic potential with a spring constant $k$ (i.e., the harmonic potential energy $U = kx^2/2$), we obtain, with the same logic,

$$\langle \frac{kx^2}{2} \rangle = \frac{1}{2} k_B T. \quad (14.52)$$
14.15 Application to homogeneous energy functions
Suppose the spatial position of a particle is governed by an anharmonic potential
\( U = k x^4 \), where \( k \) is a positive constant, then we can compute the equilibrium
average of this potential energy as
\[
\left< x \frac{\partial U}{\partial x} \right> = 4 \langle U \rangle = k_B T. \tag{14.53}
\]

Since the classical kinetic energy \( k \) is quadratic in (angular) momenta,
\[
\sum_i p_i \frac{\partial K}{\partial p_i} = 2K. \tag{14.54}
\]
Thus, if there are \( N \) particles, then there are \( 3N \) variables, so
\[
\langle K \rangle = \frac{3N}{2} k_B T. \tag{14.55}
\]
If a system is described as coupled harmonic oscillators, then the potential energy \( U \)
is a quadratic function of the position (displacement) coordinates. Therefore, quite
analogously the average total potential energy is
\[
\langle U \rangle = \frac{n_v}{2} k_B T, \tag{14.56}
\]
if there are \( n_v \) modes.

14.16 Specific heat of gases, computed classically
A direct application of the equipartition of energy is the high temperature (constant
volume) specific heat per particle of a multiatomic molecular ideal gas. Let us assume
that each molecule contains \( M \) atoms. The Hamiltonian of each molecule can be
written as
\[
H = K_{CM} + K_{rot} + K_{vib} + U_{vib}, \tag{14.57}
\]
\(^{200}\)If \( K = \sum_{i,j} A_{ij} p_i p_j \), where \( A_{ij} \) is a constant, then
\[
\sum_k p_k \frac{\partial}{\partial p_k} K = \sum_k p_k \sum_{i,j} A_{ij} (\delta_{ik} p_j + p_i \delta_{kj}) = \sum_k p_k \left( \sum_j A_{kj} p_j + \sum_i A_{ik} p_i \right) = 2K.
\]
This is an example of Euler’s theorem about homogeneous functions.
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. Notice that the total number of (angular) momenta is always $3M$ for a $M$-atomic molecule, so, obviously

$$
\langle K_{CM} + K_{rot} + K_{vib} \rangle = \frac{3}{2} Mk_B T. \quad (14.58)
$$

Thus, we have only to count the number of vibrational modes.

For a not-linear molecule there are 3 translational degrees, and 3 rotational degrees, so there are $3M - 6$ harmonic modes. Thus, $\langle U_{vib} \rangle = (3M - 6)k_B T/2$. That is, the internal energy is $E = (3M - 3)k_B T$ per molecule, so $C_V = (3M - 3)R$ per mole, where $R$ is the gas constant.

For a molecule whose shape is linear there are 3 translational degrees, and 2 rotational degrees, so there are $3M - 5$ harmonic modes. Thus, $\langle U_{vib} \rangle = (3M - 5)k_B T/2$. That is, $E = (3M - 5/2)k_B T$ per molecule, so $C_V = (3M - 5/2)R$ per mole.

It is a well-known fact that these specific heat values grossly overestimate the actual specific heats of molecular gases and were regarded as a paradox before the advent of quantum mechanics.

For a diatomic gas $M = 2$, so $C_V = (7/2)R$ according to our formula just derived, but actually around the room temperature it is usually $(5/2)R$. That is, it is less by $R$. This is because the vibrational mode is frozen and its contribution to kinetic and potential energies $R/2 + R/2$ does not show up. To excite vibration the heat bath must pay a big sum of energy (= vibrational energy quantum) at once to the molecule, so if its temperature is low, the heat bath cannot afford it. In classical mechanics the environment is allowed to pay the big sum by ‘monthly installment,’ so vibration could be excited, but in the real quantized world, this is impossible. Thus, the specific heat becomes small.

---

201 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.
Q14-1. There is a 1 mole of ideal gas consisting of molecules with one internal degree of freedom. The internal motion of an individual molecule is described by the following Hamiltonian

\[ H_{\text{int}} = \frac{1}{2} \mu p^2 + \frac{1}{4} \alpha q^4, \tag{14.59} \]

where \( \mu \) and \( \alpha \) are positive constants, and \( p \) and \( q \) are canonical coordinates describing the internal motion. The total Hamiltonian of the whole gas must be the sum of the Hamiltonian governing the center of mass translational motions of individual molecules and the Hamiltonians describing their internal motions (i.e., (14.59) for each molecule).

1. Let \( z_i \) be

\[ z_i = \frac{1}{\hbar} \int dp \int dq \, e^{-\beta H_{\text{int}}}. \tag{14.60} \]

Write down (you can copy anything usable from the lecture notes) the canonical partition function \( Z \) for this ideal gas utilizing \( z_i \). Let us assume the temperature of the gas to be \( T \), its volume \( V \) and the mass of each particle \( m \).

2. What is the constant volume specific heat of this system? [Hint: try to calculate the average of the total Hamiltonian to obtain the internal energy \( E \).]

3. Although it is possible to analytically evaluate (14.60), since we take the logarithm of \( z_i \), we have only to obtain the exponent \( \theta \) in \( z_i \propto T^\theta \). Get \( \theta \) dimensional analytically, and confirm that your result agrees with (or is consistent with) (2). [Hint. Find the dimension of \( \beta \mu \), etc.]

4. What is the constant pressure specific heat \( C_P \) of the system?

5. Classically \( C_V \) does not depend on \( \alpha \), but quantum mechanically it is not the case. Suppose the temperature goes very close to \( T = 0 \) (or \( \alpha \) becomes extremely large), what do you expect to happen to \( C_V \)?

Soln.

1. \( Z = Z_{\text{ideal}}z_j^N \). That is,

\[ Z = \frac{1}{N!} \left( \frac{2\pi m k_B T}{\hbar^2} \right)^{3N/2} z_j^N. \]

2. The contribution of the translational motion is \( 3RT/2 \). The contribution of \( p \) is just another kinetic energy, so \( RT/2 \). The contribution of \( q \) can be obtained with the aid of the equipartition of energy:

\[ \langle \alpha q \frac{\partial H}{\partial q} \rangle = \langle \alpha q^4 \rangle = k_B T, \]

so \( RT/4 \) is its contribution. Combining all of them, we get

\[ \langle H \rangle = \frac{3}{2} RT + \frac{1}{2} RT + \frac{1}{4} RT = \frac{9}{4} RT. \]
Hence, $C_V = 9R/4$.

(3) The integral wrt $p$ has the dimension of $p$, and must be a function of $\beta \mu$. $[\beta \mu p^2] = 1$, so $[p] = [\beta \mu]^{-1/2}$. Analogously, $[q] = [\beta \alpha]^{-1/4}$. Therefore,

$$[hz_i] = [\beta \mu]^{-1/2}[\beta \alpha]^{-1/4}, \quad \Rightarrow \quad z_i \propto T^{3/4}.$$ 

Therefore, $3R/4$ is the contribution of the internal degree of freedom. Consistent.


(5) The internal motion is a kind of oscillation, and obviously there is a finite energy gap. Therefore, at lower temperatures, energy quantization makes excitation harder. Eventually, $C_V$ goes to the value without the contribution of the internal degree of freedom. That is, $3R/2$. 

224
15 Information and entropy

Summary
* The Gibbs-Shannon formula for entropy/information is explained.
* Entropy quantifies how much amount of knowledge (information, measured in terms of the number of Yes-No questions) you need to specify an individual elementary event (= microstate) of a system.

Key words
(Gibbs-)Shannon formula, information, bit, surprisal

What you should be able to do
* Explain why the Shannon formula is plausible (perhaps in terms of surprisal).

15.1 Gibbs-Shannon formula of entropy
Using the canonical formalism, let us compute entropy explicitly:

\[
TS = E - A = k_B T \log Z - k_B T \text{Tr} \left( \frac{e^{-\beta H}}{Z} \right) \log e^{-\beta H} = -k_B T \text{Tr} \left( \frac{e^{-\beta H}}{Z} \right) \log \frac{e^{-\beta H}}{Z}.
\]

That is,

\[
S = -k_B T \rho \log \rho,
\]

where \( \rho = e^{-\beta H} / Z \) is the canonical density operator, or, similarly, classically

\[
S = -k_B \int d\Gamma p \log p,
\]

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 (15.5) is often called Shannon’s formula. It is a convenient occasion to see why such a formula is a good measure of information. Shannon did not ask what information was, but
tried to quantify it.\footnote{\textit{Textbook of information theory}} The best textbook of information theory (in English) is T. M. Cover and J. A. Thomas, \textit{Elements of Information Theory} (Wiley, 1991).

\subsection*{15.2 How to quantify information}

Let $\eta(m)$ be the ‘information’ per letter we can send with a message (letter sequence) that is composed of $m$ distinct symbols. Here, the word ‘information’ should be understood intuitively. Let us assume that all the symbols are used evenly. Then, $\eta(m)$ must be an increasing function of $m$; if you are allowed to use only two symbols, 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 you can use 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 of symbols consisting of $n$ symbols. 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

$$\eta(mn) = \eta(m) + \eta(n).$$  \hfill (15.6)

Since $\eta$ is an increasing function, we conclude

$$\eta(n) = c \log n,$$  \hfill (15.7)

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 \textit{nat}; if we choose $c = 1/\log 2$ (i.e., $\eta(n) = \log_2 n$), in \textit{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 (15.7)? We can write $\eta(n) = - \log_2 (1/n)$ bits; $1/n$ is the probability for a particular symbol. $- \log_2 (1/n)$ may be interpreted as the expectation value of $- \log_2$ (probability of a symbol). This suggests that for the case with not-equal-probability occurrence of $n$ symbols with
probabilities \( \{p_1, \ldots, p_n\} \), the expectation value of the information carried by the \( i \)-th symbol should be defined as \(- \log_2 p_i \) bits. Then, the average information in bits carried by a single symbol should be defined by

\[
H(\{p_i\}) = - \sum_{i=1}^{n} p_i \log_2 p_i.
\]

(15.8)

This is called the *Shannon information formula*.\(^{204}\) When Shannon arrived at (15.8), 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.

15.3 Average surprisal

The quantity \(- \log_2 p_i \) that appears in the above is sometimes called the *surprisal* of symbol \( i \), because it measures how much we are surprised by encountering this symbol (smaller \( p \) should give more surprise). It may be easier to use the axioms for surprisal to understand the Shannon formula (15.8). The ‘extent of surprise’ \( f(p) \) we get, spotting a symbol that occurs with probability \( p \) or knowing that an event actually happens whose expected probability is \( p \), should be

1. A monotone decreasing function of \( p \) (smaller \( p \) should give us bigger surprise).
2. Nonnegative.
3. Additive: \( f(pq) = f(p) + f(q) \).\(^{205}\)

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.

15.4 Entropy vs Information

Now, we have learned two pieces as to the relation between entropy and information:

(i) \( \Delta S \) due to a process in a macrosystem is related to (when \( \Delta S > 0 \)) the number of extra YES-NO questions we need to determine the microstate of the macrosystem as accurately as before the process. For an ideal gas all the molecules are independent, so we must determine all the states of particles individually.\(^{206}\) Therefore, we must multiply the number of particles to get the number of questions to pinpoint a particular microstate of the gas.

---

\(^{204}\) 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).

\(^{205}\) We could invoke the Weber-Fechner law.

\(^{206}\) Let us not consider the indistinguishability here, for simplicity.
The Gibbs-Shannon formula: the expression of entropy of a system and the expression of the information carried by a collection of letters are identical.

Combining these two, we may conclude that entropy is the amount of knowledge/information required to pinpoint the (micro)state (= elementary event) of a system. If the volume of an ideal gas is doubled, then $\Delta S = R \log 2$ per mole. Suppose we can know the state of each molecule (within a specified error) before doubling the volume. After doubling, if we know whether each molecule is in the left or in the right half (i.e., a 1 bit/molecule information), then we can know the state of each molecule even after the volume doubling (within the same specified error). The extra knowledge per molecule required is one bit. This must correspond to the entropy increase just mentioned. Therefore, $R \log 2$ of entropy should be identical to the $N_A$ bits of information: $R \log 2 = N_A$ bits. This was already alluded in Lecture 10.

Let us look at particle configurations (maybe spin configurations) on a lattice in Fig. 15.1.

![Figure 15.1: Left: ordered; Right: a representative of a collection of disordered microstates with $N (= 9)$ microscopic entities.](image)

If a microstate is ordered (Fig. 15.1 Left), since in this idealized example, there is only two microstates, to pinpoint a microstate, we need one question: Are all particles green? How about the disordered state on the Right? In this case a macrostate corresponds to many distinct microstates. To single-out any particular microstate, we need $N$ YES-NO questions. That is, we need $N (= 9$ in this example) bits of information to describe it.\(^{207}\)

As discussed $S$ may be understood as the required (expectation) amount of information we need to pinpoint any microstate in the macrostate. Therefore, we can interpret $S$ as the measure of disorder of the macrostate seen microscopically.

\(^{207}\)Even if the pattern is disordred, if a macrostate corresponds to one of the disordered configurations such as the one in the figure, then there is one YES-NO question to determine its microstate: Is the configuration such as: the first red, the second green, the third green, etc.? You may well say that no one can conceive such a question. You are very likely to be correct, but our question is whether there is such a question or not; notice that if you did not know the state is ordered, you could not make a single question even in the Left case of Fig. 15.1.

However, your objection highlights the distinction between the disordered states and the disordered systems (such as glasses).
15.5 ‘Thermodynamic unit’ of information
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 due to a reaction involving small molecules is usually the order of a few eu. This is a reasonable value.

15.6 How to quantify the amount of knowledge (gained)
How much information do we need to know the outcome of a fair dice? We guess it is \(\log_2 6\) bits. This is the entropy of a state of a dice. Suppose you are told that the face value is larger than or equal to 3. How much information does this statement carry? Information is something that can reduce our extent of ignorance. After hearing this message, we know 4 faces are still possible, so we need 2 more YN questions to remove uncertainty completely (i.e., to get a particular elementary event). Therefore, \(\log_2 6 - \log_2 4 = \log_2 \left(\frac{3}{2}\right) > 0\) must be the information carried by the message.

Example [Information carried by messages]
(1) Suppose a positive integer is given. It must begin with one of 1, 2, ···, 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 remaining 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 this 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.
15.7 Statistical mechanics from information theory?
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’ \( p \) by maximizing the Gibbs-Shannon 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 the normalization condition):

\[
-k_B \int p \log p \, d\Gamma - \beta \int p \mathcal{H} \, d\Gamma - \lambda \int p \, d\Gamma.
\]  

(15.9)

This indeed gives

\[
p \propto e^{-\beta \mathcal{H}}.
\]  

(15.10)

The Shannon formula is derived logically from almost inescapable requirements about ‘knowing something.’ Therefore, 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.\(^{208}\) This is the so-called information-theoretical derivation of statistical mechanics.

15.8 Don’t be fooled by Jaynes
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 incorporated.

Furthermore, if we look at (15.5), we should realize that something is wrong. \( p \) there is not probability but probability density, so it is not invariant under coordinate transformation. The description, for example, 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\textsuperscript{209}

\[
S = -k_B \int d\Gamma p \log \frac{p}{q}.
\]

(15.11)

We cannot do anything without fixing \(q\). To determine it, we need a certain statistical principle.\textsuperscript{210}

15.9 What is 1/2 question?

You might wonder what 0.5 yes-no questions imply. How can we ask such a question? Suppose there are 1 red ball and 999 white balls. How many questions do you need to determine the colors of all the balls? The total entropy is in this case 11.4 bits. That is, you need 0.01 Yes-No questions to determine the color of a single ball. Let us assume that initially you only know that there are red and white balls only. First, we divide the balls into two 500 ball sets and ask if one set you choose is with all the same color or not. If the answer is yes, this single bit question determines the color of 500 balls at once. Thus, it should not be so difficult to understand what a fraction of a question means.

15.10 Summary of information vs entropy

We may safely conclude that the amount of (the average) information required to pinpoint an elementary event in the sample set \(\Omega\) is proportional to

\[
\eta = -\sum_{\omega \in \Omega} p(\omega) \log p(\omega).
\]

(15.12)

If you use \(\log_2\) (i.e., \((1/\log 2)\log\)), this is the information in bits (# of average YES-NO questions you must ask). If you use \(k_B \log\), you measure information in energy unit (per molecule). The entropy \(S\) in bits is, according to Boltzmann’s principle, given by \(\log w(E, X)\). This is the number of yes-no questions you must ask to single out a microstate that is consistent with the macrostate \((E, X)\). Increasing entropy

\textsuperscript{209}This is the (negative) Kullback-Leibler entropy, whose natural implication is supplied by large deviation theory.

\textsuperscript{210}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.

231
(ΔS > 0) implies that to pinpoint a microstate you must ask extra questions corresponding to ΔS; you need ΔS extra information. This implies that the randomness of the system has increased.
Q15-1. Suppose a student cheats in a yes-no quiz by copying the answer of an all A student who is correct with probability 95%.

(1) Assuming that the student has no idea about the correct answer, what is the amount of information (its expected value) he can gain by copying the all A student’s solution?

(2) Suppose the test is known to give more YES solutions than NO solutions (3 to 2 ratio). What is the information (per question) he gains from cheating?

Soln.

(1) Initially, this student has no information about the solution at all, so his knowledge state is 0, or his ignorance as to this problem (that can be answered by one YN question) is \( H_i = 1 \) bit (i.e., with 1 bit of information his knowledge level becomes perfect, or the initial 1 bit uncertainty becomes 0). After copying the answer his remaining uncertainty (the extent of ignorance) is the same as that of the A student. That is, \( H(0.95) \):

\[
H_f = H(0.95) = -(0.95 \log_2 0.95 + 0.05 \log_2 0.05) = 0.286 \text{ bits.}
\]

Therefore, his extent of ignorance is reduced by \( 1 - 0.286 = 0.714 \) bits. This must be the information he got from cheating.

(2) Initial entropy (i.e., the initial level of ignorance) is

\[
H_i = -(3/5) \log_2(3/5) - (2/5) \log_2(2/5) = 0.971.
\]

\( H_f \) is the same as in (1), so 0.685 bits is the gain.

The solutions of (1) and (2) may be summarized as Fig. 15.2.

Q15-2. The boiling temperature of acetic acid under 1 atm is 391 K, and the evaporation heat (= latent heat of evaporation) is about 23.7 kJ/mol.

(1) What is the entropy increase due to evaporation?

(2) Roughly, how many yes-no questions do you have to ask to specify the (single) molecular state in the gas phase as accurately as in the liquid phase?

(3) The evaporation entropy of ethanol is about 110 J/K·mol. You should have realized a big difference between this value and the value you obtained in (1). This is said to be due to dimerization: acetic acid gas (around the boiling point) consists of dimers \((\text{CH}_3\text{COOH})_2\) (due to strong hydrogen bonding, but ethanol does not make dimers in the gas phase).\(^{211}\) Is the entropy difference roughly consistent with this explanation (or not)? Give your opinion with your supporting argument.

\(^{211}\)Precisely speaking, there are also tetramers, and the average acetic acid molecules in a single gas particle seems about \( 105/60 \simeq 1.75 \).
Figure 15.2: Extent of knowledge and extent of ignorance (or the disorder of the ‘brain’). BLUE bars indicate the initial state of the would-be cheater in part (1); RED bars indicate the final state of the cheater. 1 bit ignorance implies, in this case, total ignorance. GREEN bars indicate the initial state of the would-be cheater in part (2). Reduction of uncertainty/ignorance (= entropy) is acquisition of knowledge (= information).

Soln.
(1) The entropy change due to evaporation is \( \Delta S = \frac{23700}{391} = 60.1 \text{ J/K·mol} \).
(2) This corresponds to \( 60.6 \times 0.17 = 10.3 \) bits/molecule. That is, we need about 10 Yes-No questions to determine the state of each molecule as precisely as we can do so in the liquid phase. The volume of the gas (under the condition we are interested in) is about 200 times as large as that of the liquid. This explains about 7 to 8 bits. Not very bad.
(3) Ethanol evaporation corresponds to almost 19 bits/molecule increase of entropy, so we may say that the number of questions required for ethanol is almost doubled. If we assume that roughly two molecules behave together, then the knowledge about one molecule tells us about one more molecule, so this is reasonable.
16 Specific heat of solid

Summary
* Quantization forbids incremental increase of energy, so quantization generally reduces specific heat.
* If you take into account the proper dispersion relation of a crystal, quantized harmonic oscillators can explain solid specific heat.

Key words
quantum harmonic oscillator, Debye model, $T^3$-law, equipartition of energy

What you should be able to do
* Explain why quantization usually reduces specific heats.
* Be able to compute the canonical partition function of the quantized harmonic oscillator.

16.1 Review: never think physics does not demand memorizing
Let’s collect the formulas you must remember (their names are not mentioned), no definitions of symbols are provided, but a pretty standard convention (adopted in these lectures) is followed. You must be able to explain what they are and how to use them:

\[ \Delta E = Q + W, \quad (16.1) \]
\[ dS \geq \frac{1}{T_e}d'Q. \quad (16.2) \]
\[ dE = TdS - PdV + \mu dN + \mathbf{B} \cdot d\mathbf{M} + \cdots. \quad (16.3) \]
\[ A = E - TS, \quad G = A + PV, \quad H = E + PV. \quad (16.4) \]
\[ dH = TdS + VdP + \mu dN + \mathbf{B} \cdot d\mathbf{M} + \cdots. \quad (16.5) \]
\[ dA = -SdT - PdV + \mu dN + \mathbf{B} \cdot d\mathbf{M} + \cdots. \quad (16.6) \]
\[ dG = -SdT + VdP + \mu dN + \mathbf{B} \cdot d\mathbf{M} + \cdots. \quad (16.7) \]
\[ \Delta A \leq W. \quad (16.8) \]
\[ S = k_B \log w(E, X). \]  
\[ A = -k_B T \log Z(T, X). \]  
\[ Z = \sum_E w(E, X) e^{-\beta E} = Tr e^{-\beta H} = \frac{1}{h^{3N} N!} \int d\Gamma e^{-\beta H}. \]  
\[ S = -k_B \int d\Gamma p \log p, \quad H(p) = -\sum_i p_i \log p_i. \]  

16.2 Ideal gas with internal degrees of freedom

If an ideal gas particle has internal degrees of freedom, the Hamiltonian of the gas consists of two parts:

\[ H = H_0 + H_i, \]  
where \( H_0 \) is the Hamiltonian of the translational motion

\[ H_0 = \sum_{i=1}^{N} \frac{1}{2m} p_i^2, \]  
and \( H_i \) is the Hamiltonian governing the internal degrees of freedom, which is a sum of Hamiltonians \( h_i \) governing individual molecular internal motions:

\[ H_i = \sum_{i=1}^{N} h_i. \]  
The translational degrees and internal degrees of freedom are not interacting, so they are completely independent (mechanically and statistically). Hence, the canonical partition function reads

\[ Z = Z_{\text{ideal}} Z_i, \]  
where \( Z_{\text{ideal}} \) is the partition function for a monatomic ideal gas we computed before, and \( Z_i \) is the “internal” partition function

\[ Z_i = z^N, \]  
with

\[ z = \sum e^{-\beta h}. \]  
Here, the suffix to denote a particular molecule has been dropped, since all the internal partition functions are identical for identical molecules.
### 16.3 Collection of harmonic oscillators: classical approach

An important internal motion of gas molecules is vibration. Let us consider a diatomic molecule whose vibrational degree of freedom may be described as a 1D harmonic oscillator of (effective mass) $m$ and angular frequency $\omega$:

$$ h = \frac{1}{2m}(p^2 + m^2 \omega^2 q^2). \quad (16.19) $$

Although we already know that this oscillator should not be treated classically, let us study it classically to know how bad the result is:

$$ z = \frac{1}{h} \int dp dq e^{-\beta(p^2 + m^2 \omega^2 q^2)/2m} = \frac{1}{h} \left( \frac{2\pi m}{\beta} \right)^{1/2} = \frac{k_B T}{\hbar \omega}. \quad (16.20) $$

Therefore,

$$ Z_i = \left( \frac{k_B T}{\hbar \omega} \right)^N = (\beta \hbar \omega)^{-N}. \quad (16.21) $$

At this juncture, we must realize that the partition function $Z_i$ is also the canonical partition function of a collection of 1D oscillators sitting individually at lattice points.

The internal energy of a collection of $N$ 1D oscillators is

$$ E = N \frac{\partial}{\partial \beta} \log(\beta \hbar \omega) = N/\beta = N k_B T. \quad (16.22) $$

Therefore,

$$ S = (E - A)/T = N k_B + N k_B \log(k_B T/\hbar \omega) = N k_B \log T + \text{const.} \quad (16.23) $$

The result implies that in the $T \to 0$ limit $S \to -\infty$. Therefore, to describe a harmonic oscillator at $T$ knowing only the state information at much lower temperature $T_0$ a large amount of additional information $\sim \log(T/T_0)$ is required, especially if $T_0$ is close to zero. This must violate a certain fundamental principle: the finiteness principle that a finite object with finite energy should require only a finite amount of information for its complete description (if totally isolated).\(^{212}\)

This is a good occasion to discuss another principle of thermodynamics: the third law.

\(^{212}\langle \langle \text{Consequence of the finiteness principle} \rangle \rangle\) If we combine Boltzmann’s principle and this finiteness principle, then Boltzmann’s principle must not be applicable to an extremely small phase volumes. That is there must be ‘information quantum’ that has the dimension of (action)\(^3N\). Thus, notice that the existence of a fundamental quantity with the dimension of action (something like $\hbar$) is required by the finiteness principle and thermodynamics.

---

212\langle \langle \text{Consequence of the finiteness principle} \rangle \rangle\) If we combine Boltzmann’s principle and this finiteness principle, then Boltzmann’s principle must not be applicable to an extremely small phase volumes. That is there must be ‘information quantum’ that has the dimension of (action)\(^3N\). Thus, notice that the existence of a fundamental quantity with the dimension of action (something like $\hbar$) is required by the finiteness principle and thermodynamics.

237
16.4 The third law of thermodynamics
Nernst empirically found that all the entropy changes asymptotically vanish in the $T \to 0$ limit. In particular, all the derivatives of entropy $S$ vanish as $T \to 0$ (Nernst’s law). All the specific heat vanishes as $T \to 0$. Nernst concluded that entropy becomes constant (independent of any thermodynamic variables) in the $T \to 0$ limit. Later, Planck chose this constant to be zero (the concept of absolute entropy).

We adopt the following as the third law of thermodynamics:

Reversible change of entropy $\Delta S$ vanishes in the $T \to 0$ limit.

This implies that to describe the state of a macroscopic system at $T = 0$, required information is subextensive, or the number of YES-NO questions needed to know the macrostate is zero per particle.

The entropy we computed classically above must be wrong.

16.5 3D crystal; classical treatment
Consider a crystal made of $N$ atoms, having $3N$ mechanical degrees of freedom. Small displacements of atoms around their mechanical equilibrium positions should be a kind of harmonic oscillation. Thus, we may regard the crystal as a set of $3N$ independent harmonic oscillators (modes) of various frequencies (due to coupling among atoms). The canonical partition function of the total system is the product of the canonical partition function for each harmonic mode.

Treating the system completely classically and using the definition of the classical partition function (16.20), the partition function reads

$$Z = \prod_{i=1}^{3N} \frac{k_B T}{\hbar \omega_i}. \quad (16.24)$$

The contribution of these oscillators to the internal energy is readily obtained from the equipartition of energy as

$$E = 3N k_B T. \quad (16.25)$$

If the volume is kept constant, the frequencies are also kept constant. Therefore, the constant volume specific heat $C_V$ is given by

$$C_V = 3N k_B. \quad (16.26)$$

This is called Dulong-Petit’s law, which is independent of temperature, a contradiction to the third law of thermodynamics: $C_V \to 0$ in the $T \to 0$ limit.
Its entropy is just as (16.23)

\[ S = 3Nk_B \log T + \text{const.} \] (16.27)

Thus, entropy goes to \(-\infty\) as \(T \to 0\). This of course contradicts the third law.

### 16.6 Necessity of quantization

We must guess that \(C_V \downarrow 0\) should be a quantum effect. Quantization of energy implies that you cannot pay the energy cost ‘in installments.’ Since \(T\) indicates your energy payment capability, if \(T\) is smaller, then quantized systems are generally harder to excite. Recall Schottky type specific heat: it goes to zero in the \(T \to 0\) limit because of the energy gap: to excite the system, you must pay this amount at once.\(^{213}\)

### 16.7 1D quantum oscillator

Consider a collection of \(3N\) 1-dimensional harmonic oscillators which are not interacting with each other at all.

Let us first examine a single oscillator of frequency \(\nu\) (angular frequency \(\omega = 2\pi \nu\)). Elementary quantum mechanics tells us that the energy of the system is quantized as

\[ \varepsilon = \left( \frac{1}{2} + n \right) \hbar \omega, \quad n = 0, 1, 2, \cdots. \] (16.28)

Each eigenstate is nondegenerate. Thus, if we specify the quantum number \(n\), the microscopic state of a single oscillator is completely specified. The canonical partition function for a single oscillator reads

\[ z = \sum_{n=0}^{\infty} \exp \left[ -\beta \left( \frac{1}{2} + n \right) \hbar \omega \right]. \] (16.29)

Using \((1 - x)^{-1} = 1 + x + x^2 + x^3 + \cdots (|x| < 1)\), we get

\[ z = e^{-\beta \hbar \omega / 2} (1 - e^{-\beta \hbar \omega})^{-1} = \left( 2 \sinh \frac{\beta \hbar \omega}{2} \right)^{-1}. \] (16.30)

\(^{213}\)At higher temperatures, the specific heat again goes to zero in this case, because there is nothing remaining to be excited; even if you are rich, now you have nothing to buy.
16.8 Einstein model of crystal

If we may understand a 3D crystal as a collection of identical $3N$ independent 1D oscillators, the canonical partition function for the system should be

$$Z = z^{3N}. \quad (16.31)$$

From (16.31) we obtain

$$A(N) = 3Nk_B T \log \left(2 \sinh \frac{\beta \hbar \omega}{2}\right), \quad (16.32)$$

and

$$E = \frac{3}{2}N\hbar \omega \coth \left(\frac{\beta \hbar \omega}{2}\right) = 3N \left(\frac{1}{2} \hbar \omega + \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}\right). \quad (16.33)$$

Hence, the specific heat is

$$C_V = 3Nk_B \left(\frac{\hbar \omega}{k_B T}\right)^2 \frac{e^{\beta \hbar \omega}}{(e^{\beta \hbar \omega} - 1)^2}. \quad (16.34)$$

At sufficiently high temperatures ($\hbar \omega/k_B T \ll 1$) quantum effects should not be important. As expected we recover the classical result (16.26):

$$C_V \to 3Nk_B. \quad (16.35)$$

For sufficiently low temperatures ($\hbar \omega/k_B T \gg 1$) (16.34) reduces to

$$C_V \simeq 3Nk_B \left(\frac{\hbar \omega}{k_B T}\right)^2 e^{-\beta \hbar \omega}. \quad (16.36)$$

Thus, $C_V$ vanishes at $T = 0$, and the third law behavior is exhibited, but notice that this is a Schottky type specific heat due to the energy gap of size $\hbar \omega$.

16.9 Real 3D crystal: Debye model

$C_V$ just obtained goes to zero exponentially fast at variance with the empirical law:

$$C_V \sim T^3. \quad (16.37)$$

\[214^*\]

$$\coth \frac{x}{2} = \frac{e^{x/2} + e^{-x/2}}{e^{x/2} - e^{-x/2}} = \frac{e^{x/2} - e^{-x/2} + 2e^{-x/2}}{e^{x/2} - e^{-x/2}} = 1 + \frac{2}{e^x - 1}$$

240
As we know from the Schottky type specific heat 13.3 it is the rule that whenever there is a finite energy gap \( \varepsilon \) between the ground and the first excited states, the specific heat behaves like \( \exp(-\beta \varepsilon) \) at low temperatures. The empirical result (16.37) implies that there is no finite energy gap in real crystals.

In a real crystal there is a distribution in vibrational frequencies (= dispersion) as can be seen from Fig. 16.1.

Not all the vibrations contribute significantly to the low temperature heat capacity of solids. Elastic vibrations in a crystal can be classified into two branches, optical and acoustic (Fig. 16.2).

Only the acoustic modes are relevant. We must study the number of acoustic modes with a given angular frequency about \( \omega \).

16.10 How to count the number of modes
As can be seen from Fig. 16.1 in a 1D direction the possible wavelengths are \( \lambda = \)
$2L, 2L/2, \cdots, 2L/N$ or in wave numbers $k = \pi/L, 2\pi/L, \cdots, N \pi/L$. We have already seen such a sequence before: the de Broglie wave length of a free particle confined in a box which we used to study the classical ideal gas. This implies we can compute the number of modes in the volume $V$ just as the number of eigenvalues as follows (we will encounter this approach repeatedly later, so you need not understand it now). Let the number of modes with angular frequency between $\omega$ and $\omega + d\omega$ be $\mathcal{D}(\omega)$. Then, we have

$$\int_0^\omega \mathcal{D}(\omega)d\omega = \frac{1}{\hbar^3} \int d\mathbf{r} \int_{p\leq p(\omega)} d\mathbf{p},$$

where $p(\omega) = \hbar k = \hbar\omega/c$ (dispersion relation), and $c$ is the sound speed. Therefore, the number of modes between $\omega$ and $\omega + d\omega$ is

$$\mathcal{D}(\omega) = \frac{V}{\hbar^3} 4\pi p(\omega)^2 \frac{dp(\omega)}{d\omega} = 4\pi V \frac{\hbar^3}{\hbar^3 c^3}\omega^2 = \frac{1}{2\pi^2} V \frac{\omega^2}{c^3},$$

(16.39)

In reality, there are one longitudinal and two transversal modes for each $\omega$, so the actual number is this times 3. In contrast to the classical gas case there are two important differences. $\mathcal{D}(\omega) \propto \omega^2$ holds only for low frequency modes where the material may be regarded as an elastic continuum body. Furthermore, the wavelength cannot be indefinitely small as seen from Fig. 16.1. Debye introduced the following approximation:

$$\mathcal{D}(\omega) = A\omega^2 \Theta(\omega_D - \omega),$$

(16.40)

where $\omega_D$ is the Debye cutoff frequency (which is a materials ‘constant’ in good approximation) and $A$ is fixed to have the total number of modes (= the total number of lattice cells $N \times 3$) correctly

$$\int_0^{\omega_D} \mathcal{D}(\omega)d\omega = 3N.$$

(16.41)

Therefore, $A = 9N/\omega_D^3$.

### 16.11 Debye model specific heat

Since we know the energy of the mode with $\omega$ (16.33), the total energy (the internal energy due to lattice vibration) is

$$E = \int d\omega \mathcal{D}(\omega) \left( \frac{1}{2} \hbar \omega + \frac{\hbar \omega}{e^{\hbar \omega/c} - 1} \right),$$

(16.42)
and from (16.34)

\[ C_V = k_B \int d\omega \mathcal{D}(\omega) \left( \frac{\hbar \omega}{k_B T} \right)^2 \frac{e^{\beta \hbar \omega}}{(e^{\beta \hbar \omega} - 1)^2}. \]  

(16.43)

Although the integration range has an upper bound \( \omega_D \), when the temperature is small, replacing this with \( \infty \) does not change the integral appreciably. Therefore, \( C_V \) behaves just as \( T^3 \) in the low temperature limit; the dimension of the integral is \( [\omega]^3 \), and we know \( [\hbar \omega / k_B T] = 1 \), so we may conclude \( [\omega]^3 \propto T^3 \) simply with the aid of dimensional analysis.
17 How to manipulate partial derivatives

Summary
* Understand thermodynamics and microscopic picture of the thermal properties of a rubber band.
* Review partial derivatives.
* The Jacobian technique may be fully utilized if you remember three elementary rules/formulas:

\[
\frac{\partial (X,Y)}{\partial (A,B)} = - \frac{\partial (X,Y)}{\partial (B,A)} = \frac{\partial (Y,X)}{\partial (B,A)} = - \frac{\partial (Y,X)}{\partial (A,B)},
\]

and Maxwell’s relation for conjugate pairs \((X,x)\) and \((Y,y)\):

\[
\frac{\partial (X,x)}{\partial (y,Y)} = 1.
\]

Key words
entropic elasticity, Maxwell’s relation, adiabatic cooling, adiabatic demagnetization.

What you should be able to do
* Practice the Jacobian technique.
* Intuitively explain rubber elasticity; be able to get various signs of partial derivatives, and to explain them intuitively.
* Explain adiabatic demagnetization.

17.1 Rubber band experiments
Let us perform a small experiment of quasistatic adiabatic processes using a rubber band. Prepare a thick rubber band (that is used to bundle, e.g., asparagus). Use your lip as a temperature sensor. Initially, putting the rubber band to your lip, you sense the room temperature (cool). Now, you hold the both ends of a small portion of the band with your hands and stretch it tightly and quickly (Fig. 17.1).

Then, feel the temperature of the stretched portion with your lip. It must be warm.
You have just demonstrated
\[ \frac{\partial T}{\partial L} \bigg|_S > 0, \]  \hspace{1cm} (17.1)
where \( L \) is the length of the stretched portion of the rubber band. The Gibbs relation for a rubber band reads
\[ dE = TdS + FdL, \]  \hspace{1cm} (17.2)
where \( F \) is the force (the component of the force parallel to the stretching direction of the force) stretching the band. Since the process is adiabatic and quasistatic, \( S \) is constant. Even if you rapidly pull the band, the maximum stretching rate you can realize is very small from the molecular point of view, so the process is (almost) quasistatic. Since the heat conduction is not a very rapid process, during quick stretch the system is virtually thermally isolated (= adiabatic). Thus, \( S \) is constant.

17.2 Polymer chain is just as kids playing hand in hand
To begin with, let us try to understand ‘microscopically’ what we have observed macroscopically. A rubber band is made of a bunch of polymer chains. Take a single chain that is wiggling due to thermal motion (Fig. 17.2a).

Figure 17.2:  

a: A schematic picture of a single polymer chain. Each arrow is called a monomer.  
b: Polymer-kid analogy. 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). The entropy of a chain 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.

Stretching the chain corresponds to increasing the space between the flags of Fig. 17.2b in the playing kid analogy. If the chain does not break, the spatial room
for moving is decreased, but since the kids must keep their entropy, the restricted
dancing motion must find substitute degrees of freedom: shaking bodies (i.e., room
in the momentum subspace). That is, the temperature of the system should go up.
This suggests that if the chain is stretched under constant $T$, entropy should go
down:

$$\left. \frac{\partial S}{\partial L} \right|_T < 0.$$  \hspace{1cm} (17.3)

Can you conclude this from what you observed (17.1)? Yes, you can, but you must
know how to manipulate partial derivatives (see toward the end of today’s lecture).

### 17.3 Freely jointed polymer chain

Before explaining powerful thermodynamic techniques, let us try to understand the
polymer chain system statistical-mechanically. For simplicity, let us consider a poly-
mer chain along the $x$-axis (Fig. 17.3).

![Figure 17.3: It is expanded in the vertical direction to avoid cluttering. Monomers can take the + or − direction. The right and left direction monomer numbers $N_\pm (N_+ + N_- = N)$ can be used to compute the number of conformations.](image)

We assume that the chain is free-jointed, that is, there is no energy cost to change
its conformation at all (just as an ideal gas can change its configuration without
any energy cost). The Hamiltonian of this free-jointed polymer consists of the chain
kinetic energy $K$ only, which is, in any case, independent of the conformation of the
chain. The Gibbs relation is

$$dE = TdS + FdL,$$  \hspace{1cm} (17.4)

where, as above, $L$ is the length (i.e., the end-to-end distance) of the chain, and $F$
is the stretching force. Do not forget that $E$ depends only on $T$, since it is (just the
averaged) $K$.\footnote{Since $E$ does not depend on $L$, you might wonder where the work we do is stored when we stretch the rubber. Since $E$ does not depend on $L$, it is surely stored not in the chain conformations, but is stored as the kinetic energy (so the temperature goes up). The situation is exactly the same as the ideal gas. In this case, the volume work done to the system is stored as its kinetic energy, since $E$ does not depend on $V$.}
We can easily compute the entropy of the ideal rubber band using Boltzmann’s principle. Following the figure caption of Fig. 17.3, let us introduce \( N_\pm \) so that \( N_+ + N_- = N \), and \( N_+ - N_- = X \equiv L/\ell \). We have

\[
N_\pm = \frac{1}{2}(N \pm X).
\]  

Therefore, (recognize that this is exactly the same problem as the Schottky defect problem as to obtaining the entropy)

\[
w(X) = \left( \frac{N}{N_+} \right). \tag{17.6}
\]

From this, immediately we obtain

\[
S = -Nk_B \left[ \frac{N_+}{N} \log \frac{N_+}{N} + \frac{N_-}{N} \log \frac{N_-}{N} \right]. \tag{17.7}
\]
or

\[
S = -Nk_B \left[ \frac{N + X}{2N} \log \frac{N + X}{2N} + \frac{N - X}{2N} \log \frac{N - X}{2N} \right]. \tag{17.8}
\]

With the aid of the Gibbs relation, we obtain (note that \( \ell X = L \))

\[
F = -\frac{T}{\ell} \frac{\partial S}{\partial X} = \frac{k_B T}{2\ell} \log \frac{N + X}{N - X}. \tag{17.9}
\]

This implies

\[
L = N\ell \tanh(\beta \ell F). \tag{17.10}
\]

Since \( \tanh x \approx x \) for small \( x \), this implies a Hookean spring

\[
F = (k_B T/N\ell^2)L. \tag{17.11}
\]

That is, \( k_B T/\langle R^2 \rangle \) is the spring constant, where \( \langle R^2 \rangle = N\ell^2 \) is the mean square end-to-end distance of a polymer chain.

### 17.4 Ideal rubber band

Can we explain what we have experienced at the beginning of this lecture using this entropy? Never. \( \text{(17.8)} \) implies that if \( L \) is fixed, then \( S \) is fixed. This is physically obvious, because the set of allowed conformations is completely determined by \( L \). It is clear that we need thermal motion. Then, the entropy should read

\[
S = Nk_B \left[ \frac{N + X}{2N} \log \frac{N + X}{2N} + \frac{N - X}{2N} \log \frac{N - X}{2N} \right] + S_e(E) \tag{17.12}
\]

247
where \( S_e \) is the entropy dependent only on the internal energy \( E \) which is solely due to thermal motion.\(^{216}\) This implies that the temperature-internal energy relation is independent of \( L \) (just as the internal energy of the ideal gas is independent of \( V \)). Such a rubber is called an ideal rubber.

We can make a more detailed model of a rubber band to compute more realistic entropy, but without such microscopic details thermodynamics can tell you many qualitative features. To this end, we must be able to manipulate thermodynamic quantities and derivatives more efficiently. Let us begin with a review of partial differentiation.

### 17.5 Partial derivative review

Let us write down the definition of partial derivatives. Consider a two-variable function \( f = f(x, y) \). Then, partial derivatives are defined as

\[
\frac{\partial f}{\partial x} \equiv f_x(x, y) = \lim_{\delta x \to 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x}, \quad (17.13)
\]

\[
\frac{\partial f}{\partial y} \equiv f_y(x, y) = \lim_{\delta y \to 0} \frac{f(x, y + \delta y) - f(x, y)}{\delta y}. \quad (17.14)
\]

Partial differentiation is extremely tricky in general. For example, even if \( \partial f/\partial x \) and \( \partial f/\partial y \) exist at a point, \( f \) can be discontinuous at the same point. \( E \) is once continuously differentiable with respect to \( S \) and work coordinates, so we need a stronger concept of differentiability in the many-variable case.

### 17.6 Derivative of multivariate function

Let \( f \) be a function of several variables \( \mathbf{x} = (x_1, \ldots, x_n) \). We could understand \( f \) as a function of the vector \( \mathbf{x} \). We wish to study its ‘linear response’ to the change \( \mathbf{x} \to \mathbf{x} + \delta \mathbf{x} \):

\[
\delta f(\mathbf{x}) = f(\mathbf{x} + \delta \mathbf{x}) - f(\mathbf{x}) = Df(\mathbf{x})d\mathbf{x} + o[\delta \mathbf{x}], \quad (17.15)
\]

where \( o \) denotes higher order terms that vanish faster than \( \|\delta \mathbf{x}\| \), when the limit \( \delta \mathbf{x} \to 0 \) is taken. Here, \( Df \) is a linear operator\(^{217}\) that can be written as

\[
Df(\mathbf{x})d\mathbf{x} = \sum_i \frac{\partial f}{\partial x_i} dx_i. \quad (17.16)
\]

\(^{216}\)Recall that the ideal gas entropy has this form.

\(^{217}\)[This is a repetition for convenience.] \( L \) is a linear operator acting on a function set \( \mathcal{F} \), if it is
If such a linear map $Df$ is well-defined, we say that $f$ is (totally) differentiable (or strongly differentiable). If there are only two variables, we may write

$$Df(x, y)(dx, dy) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$  \hfill (17.17)

### 17.7 Maxwell’s relations

Let us closely look at $f(x + \delta x, y + \delta y) - f(x, y)$. There are two ways to go from $(x, y)$ to $(x + \delta x, y + \delta y)$, $\delta x$ first or $\delta y$ first:

1. $f(x + \delta x, y + \delta y) - f(x, y) = f(x + \delta x, y + \delta y) - f(x + \delta x, y) + f(x + \delta x, y) - f(x, y) = f_y(x + \delta x, y) \delta y + f_x(x, y) \delta x$, \hfill (17.18)
2. $f(x + \delta x, y + \delta y) - f(x, y) = f(x + \delta x, y + \delta y) - f(x, y + \delta y) + f(x, y + \delta y) - f(x, y) = f_x(x, y + \delta y) \delta x + f_y(x, y) \delta y$. \hfill (17.19)

The difference between these two formulas is

$[f_y(x + \delta x, y) - f_y(x, y)] \delta y - [f_x(x, y + \delta y) - f_x(x, y)] \delta x = [f_{xy}(x, y) - f_{yx}(x, y)] \delta x \delta y$. \hfill (17.20)

This must vanish if the surface defined by $f$ is at least twice differentiable. That is,

$$f_{xy} = f_{yx}. \hfill (17.21)$$

For example, for a rubber band $dE = TdS + FdL$, so

$$\frac{\partial T}{\partial L} \bigg|_S = \frac{\partial F}{\partial S} \bigg|_L. \hfill (17.22)$$

Such relations are called Maxwell’s relations in thermodynamics.

### 17.8 Jacobian technique

To manipulate many partial derivatives, it is very convenient to use the so-called a map from $\mathcal{F}$ to some other vector space such that

$$L(\alpha f + \beta g) = \alpha Lf + \beta Lg,$$

where $f, g \in \mathcal{F}$ and $\alpha, \beta$ are numbers.

\textsuperscript{218}We must say something more careful mathematically, but let us be contented with this for now.
**Jacobian technique.** This technique may not even be taught in graduate courses, but it is easy to memorize, and easy to use. It can greatly reduce the insight and skill required in thermodynamics, especially with the Jacobian version of Maxwell’s relation (17.39).\(^{219}\)

The Jacobian for two functions \(X\) and \(Y\) of two independent variables \(x, y\) is defined by the following determinant:

\[
\frac{\partial (X, Y)}{\partial (x, y)} \equiv \begin{vmatrix}
\frac{\partial X}{\partial x} & \frac{\partial X}{\partial y} \\
\frac{\partial Y}{\partial x} & \frac{\partial Y}{\partial y}
\end{vmatrix} = \frac{\partial X}{\partial x} \frac{\partial Y}{\partial y} - \frac{\partial Y}{\partial x} \frac{\partial X}{\partial y}.
\] (17.23)

In particular, we observe

\[
\frac{\partial (X, y)}{\partial (x, y)} = \frac{\partial X}{\partial x} \Bigg|_y,
\] (17.24)

which is the key observation of this technique. Obviously,\(^{220}\)

\[
\frac{\partial (X, Y)}{\partial (X, Y)} = 1.
\] (17.25)

There are only two or three formulas you should learn by heart (they are very easy to memorize). One is straightforwardly obtained from the definition of determinants: exchanging rows or columns switches the sign:

\[
\frac{\partial (X, Y)}{\partial (x, y)} = -\frac{\partial (X, Y)}{\partial (y, x)} = \frac{\partial (Y, X)}{\partial (y, x)} = -\frac{\partial (Y, X)}{\partial (x, y)}.
\] (17.26)

Also notice that for a constant \(c\)

\[
\frac{\partial (cX, Y)}{\partial (x, y)} = c \frac{\partial (X, Y)}{\partial (x, y)}.
\] (17.27)

**17.9 Chain rule in terms of Jacobians**

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:

\[
\frac{\partial (X, Y)}{\partial (a, b)} \frac{\partial (a, b)}{\partial (x, y)} = \frac{\partial (X, Y)}{\partial (x, y)}.
\] (17.28)

\(^{219}\)Mathematically rigorously speaking, some people would claim that there can be many dangerous things, but in thermodynamics, the manipulation is perfectly mechanical (formal), so virtually there is no danger of making any logical errors.

\(^{220}\)In this case, we regard \(X\) and \(Y\) are independent variables. In the Jacobian expression, all the letters appearing upstairs are regarded dependent variables of the variables itemized downstairs.
This is a disguised chain rule:

\[
\frac{\partial X}{\partial x} \bigg|_y = \frac{\partial X}{\partial a} \bigg|_b \frac{\partial a}{\partial x} \bigg|_y + \frac{\partial X}{\partial b} \bigg|_a \frac{\partial b}{\partial x} \bigg|_y,
\]

(17.29)

e tc. Confirm (17.28) by yourself (use \(det(AB) = (detA)(detB)\)).

The technical significance of (17.28) must be obvious; calculus becomes algebra! You may think \(\partial(A, B)\) just as an ordinary number: formally we can do as follows.\(^{221}\)

First split the ‘fraction’ and then throw in the identical factors you wish to introduce:

\[
\frac{\partial(X, Y)}{\partial(x, y)} = \frac{\partial(X, Y)}{\partial(x, y)} \frac{\partial(A, B)}{\partial(x, y)} = \frac{\partial(X, Y)}{\partial(A, B)} \frac{\partial(A, B)}{\partial(x, y)}.
\]

(17.30)

From (17.28) we get at once

\[
\frac{\partial(X, Y)}{\partial(A, B)} = 1 \left/ \frac{\partial(X, Y)}{\partial(A, B)} \right.
\]

(17.31)

In particular, we have\(^{222}\)

\[
\frac{\partial X}{\partial x} \bigg|_Y = 1 \left/ \frac{\partial X}{\partial Y} \right.
\]

(17.32)

17.10 Some illustrations of Jacobian technique

Using these relations, we can easily demonstrate

\[
\frac{\partial X}{\partial y} \bigg|_x = - \frac{\partial X}{\partial y} \bigg|_x \left/ \frac{\partial X}{\partial X} \bigg|_y \right.
\]

(17.33)

as follows:

\[
\frac{\partial(X, x)}{\partial(y, x)} = \frac{\partial(X, x)}{\partial(y, X)} \frac{\partial(y, X)}{\partial(y, x)} \frac{\partial(y, x)}{\partial(X, y)} = \frac{\partial(x, X)}{\partial(y, X)} \frac{\partial(X, y)}{\partial(x, y)}.
\]

(17.34)

Then, use (17.31). A concrete example of this formula is

\[
\frac{\partial P}{\partial T} \bigg|_V = - \frac{\partial V}{\partial T} \bigg|_P \left/ \frac{\partial V}{\partial P} \bigg|_T \right.
\]

(17.35)

\(^{221}\)In pragmatic thermodynamics, you can be maximally formal and seldom make any mistake.

\(^{222}\)Mathematically properly speaking, on the LHS we regard \(X\) and \(Y\) as functions of \(x\) and \(y\), and \(Y = Y(x, y)\) is fixed. In contrast, on the RHS \(x\) and \(y\) are understood as a function of \(X\) and \(Y\), and \(Y\) is being kept constant.
which relates thermal expansivity and isothermal compressibility.

For a rubber band

\[
\frac{\partial L}{\partial S} \bigg|_F = \frac{\partial (L, F)}{\partial (S, F)} = \frac{\partial (L, F)}{\partial (T, F)} \frac{\partial (T, F)}{\partial (S, F)} = \frac{\partial L}{\partial T} \bigg|_F \frac{\partial T}{\partial S} \bigg|_F,
\]

which reads

\[
\frac{\partial L}{\partial S} \bigg|_F = \frac{\partial L}{\partial T} \bigg|_F \frac{\partial T}{\partial S} \bigg|_F = T \frac{\partial L}{\partial T} \bigg|_F \frac{1}{C_F}.
\]

Here, \(C_F\) is the heat capacity under constant force. It is explained in 17.11.

17.11 Specific heat and entropy

The relation between heat and entropy (Clausius’ equality) tells us \(dQ = TdS\), so if we differentiate this with respect to \(T\) under constant \(F\), it must be the heat capacity under constant \(F\). Generally speaking, the heat capacity under constant \(X\) (which can be extensive or intensive) always have the following expression:

\[
C_X = T \frac{\partial S}{\partial T} \bigg|_X.
\]

We will learn the consequence of the stability of the equilibrium states later, but the stability of the equilibrium state implies \(C_X \geq 0\) (usually strictly positive). Imagine the contrary. If you inject heat into a system, its temperature goes down, so it sucks more heat from the surrounding world, and further reduces its temperature. That is, such a system becomes a bottomless heat sink.

17.12 (Unified) Maxwell’s relation

All the Maxwell’s relations can be unified in the following form

\[
\frac{\partial (X, x)}{\partial (Y, y)} = -1,
\]

where \((x, X)\) and \((y, Y)\) are conjugate pairs. This is the third equality you should memorize. When you use this, do not forget that \((-P, V)\) (not \((P, V)\)) is the conjugate pair.

Let us demonstrate this. From \(\cdots + x dX + y dY + \cdots\) Maxwell’s relation reads

\[
\frac{\partial x}{\partial Y} \bigg|_X = \frac{\partial y}{\partial X} \bigg|_Y.
\]
That is,
\[ \frac{\partial(x, X)}{\partial(Y, X)} = \frac{\partial(y, Y)}{\partial(X, Y)}. \] (17.41)

This implies (mere \( a/b = c/d \Rightarrow a/c = b/d \))
\[ \frac{\partial(x, X)}{\partial(y, Y)} = \frac{\partial(Y, X)}{\partial(X, Y)} = -1. \] (17.42)

Perhaps, the following may be better:\textsuperscript{223}
\[ \frac{\partial(X, x)}{\partial(y, Y)} = 1. \] (17.43)

For example, (17.22) can be obtained as follows:
\[ \frac{\partial F}{\partial T} \bigg \rvert_{\mathcal{S}} = \frac{\partial(T, S)}{\partial(L, S)} = \frac{\partial(L, F)}{\partial(L, S)} \frac{\partial(T, S)}{\partial(L, F)} = \frac{\partial F}{\partial S} \bigg \rvert_{\mathcal{L}}. \] (17.44)

\section{17.13 Rubber band thermodynamics}
Equipped with the machinery, let us study the rubber band in more detail. The rubber band is elastic because of the thermal motion of the polymer chains. That is, resistance to reducing entropy is the cause of elastic bouncing. Thus, such elasticity is called the \textit{entropic elasticity}.\textsuperscript{224} An important feature is that the elastic force increases with \( T \) under constant length (which is easily understood from the kid picture Fig. 17.2):
\[ \frac{\partial F}{\partial T} \bigg \rvert_{\mathcal{L}} > 0. \] (17.45)

Is this related to what we have observed (17.1)? Yes. Follow the following logic (as a practice):
\[ 0 < \frac{\partial T}{\partial L} \bigg \rvert_{\mathcal{S}} = \frac{\partial(T, S)}{\partial(L, S)} = \frac{\partial(T, S)}{\partial(L, F)} \frac{\partial(L, F)}{\partial(L, S)} = \frac{\partial(L, F)}{\partial(T, L)} \frac{\partial(T, S)}{\partial(L, S)} \frac{\partial(F, L)}{\partial(T, L)} / \frac{\partial(S, L)}{\partial(T, L)} = \frac{\partial F}{\partial T} \bigg \rvert_{\mathcal{L}} \frac{T}{C_L}. \] (17.46)

\textsuperscript{223}If you know the theory of differential forms, this must be trivial: since \( d^2 E = d(\cdots + x dX + y dY + \cdots) = 0 \), if you change only \( X \) and \( Y \), \( dx \wedge dX = -dy \wedge dY \). The ratio of the infinitesimal areas on both sides is the Jacobian. Thus, the relation is not due to the conservation law, but due to thermodynamic quantities being state variables.

\textsuperscript{224}In contrast, the usual elasticity is called \textit{energetic elasticity}, which is caused by opposing increase of energy.
Or

\[ 0 < \frac{\partial T}{\partial L}
\]

\[ \middle|_S = \frac{\partial(T, S)}{\partial(L, S)} = \frac{\partial(T, S)}{\partial(T, L)} \frac{\partial(T, L)}{\partial(L, S)} = -\frac{\partial(T, S)}{\partial(T, L)} \frac{T}{C_L} \]  \hspace{1cm} (17.48)

\[ = -\frac{\partial(F, L)}{\partial(T, L)} \frac{\partial(T, S)}{\partial(F, L)} \frac{T}{C_L} = \frac{\partial F}{\partial T} \bigg|_L \frac{T}{C_L}. \]  \hspace{1cm} (17.49)

That is, you do not need any foresight.

From our microscopic imagination visualized in Fig. 17.2, we guessed

\[ \frac{\partial S}{\partial L} \bigg|_T < 0. \]  \hspace{1cm} (17.50)

Let us derive this from (17.1).

\[ \frac{\partial S}{\partial L} \bigg|_T = \frac{\partial(S, T)}{\partial(L, T)} = \frac{\partial(S, T)}{\partial(L, S)} \frac{\partial(L, S)}{\partial(L, T)} = -\frac{\partial T}{\partial L} \bigg|_S \frac{C_L}{T} < 0. \]  \hspace{1cm} (17.51)

17.14 Adiabatic cooling with rubber band

If a tightly stretched rubber band is suddenly relaxed (= adiabatically relaxed) after equilibrating with the room temperature, what do you observe? The band becomes very cool. This is not surprising:

\[ \frac{\partial T}{\partial L} \bigg|_S > 0; \]  \hspace{1cm} (17.52)

Now, \( L \) is reduced under constant \( S \), so must decrease \( T \). This is the principle of adiabatic cooling (see Fig. 17.4).

17.15 Cooling via adiabatic demagnetization

Unfortunately, we cannot use a rubber band to cool a system to a very low temperature, since it becomes brittle (chain motion freezes out easily). In actual low temperature experiments, a collection of almost non-interacting magnetic dipoles (i.e., a paramagnetic material) is used. The system is closely related to polymer chains as illustrated in Fig. 17.5.
Figure 17.4: Initially, the system is at $T_1$. Isothermally, $L$ is increased as $L_1 \rightarrow L_2$. This decreases entropy. Now, $L$ is returned to the original smaller value adiabatically and reversibly. The entropy is maintained, and the temperature decreases (adiabatic cooling) to $T_2$. The dotted path is the one you experienced by initial rapid stretching of a rubber band.

Figure 17.5: a of Fig. 17.2 corresponds to b a paramagnet, a collection of only weakly interacting magnetic dipoles.

The Gibbs relation of the magnetic system is

$$dE = TdS + BdM,$$  \hspace{1cm} (17.53)

where $B$ is the magnetic field, and $M$ the magnetization. The correspondences $B \leftrightarrow F$ and $M \leftrightarrow L$ are almost perfect: $M$ is the sum of small dipole vectors, and $L$ is also the sum of (the projected components) of ‘steps’ (monomer orientation vectors). Thus, we expect

$$\left. \frac{\partial T}{\partial M} \right|_S > 0$$  \hspace{1cm} (17.54)

and adiabatic cooling can be realized; first apply a strong magnetic field and align all the dipoles. We can do this slowly and isothermally. Then, turn off the magnetic field to make $M \rightarrow 0$ (demagnetization). Simply replacing $L$ with $M$ in Fig. 17.4, we can understand this adiabatic demagnetization strategy to cool a system.

17.16 Ideal magnetic system

We can imagine a collection of noninteracting magnetic dipoles (called spins) taking only up or down (or $s = \pm \mu$) values.\(^{225}\) The total magnetization of the system reads

$$M = \mu \sum_{i=1}^{N} s_i,$$  \hspace{1cm} (17.55)

\(^{225}\)If magnetic atoms are dilute in an insulating solid, they do not interact with each other.
where \( \mu \) is the ratio of the magnetic moment and the spin (the gyromagnetic ratio). It is a good exercise to compute \( S \) as a function of \( M \), but, as you can guess easily,

\[
M = \mu N \tanh \beta \mu B,
\]  
(17.56)

where \( B \) is the magnetic field. The relation between \( M \) and \( B \) for small \( B \) corresponds to Hooke’s law (17.11):

\[
M = (N\mu^2/k_B T)B,
\]  
(17.57)

which is called Curie’s law: the magnetic susceptibility \( \chi = N\mu^2/k_B T \).

Just as in the freely-jointed polymer model we discussed (without the \( S_e \) term), there is no kinetic energy of spins (or of the entities carrying spins), so the entropy of this model is constant under constant \( M \). Thus, just as in the ideal rubber model, without the term similar to \( S_e \) the model cannot explain the use of adiabatic demagnetization to cool other systems as a refrigerating mechanism. However, if we are interested in the spins themselves, then their coupling to other degrees of freedom (the so-called spin-lattice coupling) should not be large. Under this condition, the system can be described by the present model, so under the adiabatic demagnetization condition \( M \) is constant, because \( S \) is constant. Thus, since (17.57) or (17.56) implies \( B/T \) is constant, reducing \( B \) implies decreasing \( T \).

If we use the magnet as a coolant to cool other systems, we are interested in

\[
\left. \frac{\partial T}{\partial B} \right|_S = \frac{\partial (T, S)}{\partial (B, S)} = \frac{\partial (T, S)}{\partial (B, M)} \frac{\partial (B, M)}{\partial (B, S)} = -\frac{\partial (B, M)}{\partial (B, S)}
\]  
(17.58)

\[
= \frac{\partial (B, T)}{\partial (B, S)} \frac{\partial (B, M)}{\partial (B, T)} = -\frac{T}{C_B} \left. \frac{\partial M}{\partial T} \right|_B.
\]  
(17.59)

Therefore, if Curie’s law of the form \( M = a(B/T) \) holds, then

\[
\delta T = \frac{aB}{C_B T} \delta B
\]  
(17.60)

gives the cooling rate.

---

\(^{226}\)For a rubber band, reducing \( F \) while keeping \( L \) constant in order to change \( T \) is experimentally unthinkable.
Q17.1 [Basic problems]

(1) \( F = x \sin y, \) and \( y = x + z. \) Express

\[
\left. \frac{\partial F}{\partial x} \right|_y \quad \text{and} \quad \left. \frac{\partial F}{\partial x} \right|_z
\]

in terms of \( x \) and \( y. \) \([\sin y; \sin y + x \cos y]\)

(2) For a gas \( PV \) and \( E \) are functions of \( T \) only. Show that actually \( PV/T \) is a constant. \([\text{Compute } (\partial E/\partial V)_T = T(\partial S/\partial V)_T - P = T(\partial P/\partial T)_V - P = 0.]\)

(3) For a general gas, find the temperature change \( dT \) due to adiabatic free expansion \( V \rightarrow V + dV. \) \([\text{Compute } (\partial T/\partial V)_E. \ (\partial T/\partial V)_E = [P - T(\partial P/\partial T)_V]/C_V]\)

(4) If \( M \) is a function of \( B/T, \) \( E \) is a function of \( T \) only. \([\text{Show } (\partial E/\partial B)_T = 0 \text{ using } (\partial B/\partial T)_M = B/T; \) quite parallel to (2).\]
18 Stability, fluctuation, and response

Summary
* We can derive the universal stability (and evolution) criterion \( \delta^2 S < 0 \) (or \( \delta^2 S > 0 \)) that is independent of the environmental constraints (say, isothermal, constant volume or not, etc.)
* \( \partial (X, Y) / \partial (x, y) > 0 \).
* In equilibrium changes occur in the direction to discourage further changes (to avoid run-away processes) (Le Chatelier-Braun principle). Our world is stable!
* A generalized Gibbs free energy: \( \tilde{G}(T, x) = -k_B T \log \sum e^{-\beta (H - x \hat{X})} \).
* Susceptibilities are directly related to the second moments of fluctuations (fluctuation-response relation).

Key words

What you should be able to do
* Be able to derive the universal stability criterion.
* To mention some of the crucial conclusions due to the stability criterion (say, \( C_X > 0 \), \( C_x > C_X \), etc.).
* Be able to build a convenient partition function to obtain a convenient thermodynamic potential directly.
* Be able to recognize important conclusions we can obtain from the fluctuation-response relation.
* Be able to explain why fluctuation studies are important.

18.1 Stability question and need for stability criteria
What is the sign of
\[
\frac{\partial S}{\partial F} \bigg|_L
\]
for a rubber band? Intuition tells us that it must be positive (To increase \( F \) while keeping \( L \), we must invigorate the motion of chains, so we must raise the temperature, resulting in the increase of entropy). Let us check this, starting with our empirical
result
\[ \frac{\partial T}{\partial L} \bigg|_s > 0. \quad (18.2) \]

What you should do first is to rewrite the partial derivative in terms of Jacobians:
\[ \frac{\partial S}{\partial F} \bigg|_L = \frac{\partial(S, L)}{\partial(F, L)}. \quad (18.3) \]

(18.2) is
\[ \frac{\partial T}{\partial L} \bigg|_s = \frac{\partial(T, S)}{\partial(L, S)}, \quad (18.4) \]

so we should keep \((L, S)\), which appears in both the formulas above, and insert \((T, S)\):
\[ \frac{\partial S}{\partial F} \bigg|_L = \frac{\partial(S, L)}{\partial(F, L)} = \frac{\partial(S, L)}{\partial(T, S)} \frac{\partial(T, S)}{\partial(F, L)} = -\frac{\partial(S, L)}{\partial(T, S)} = \frac{\partial(L, S)}{\partial(T, S)} > 0. \quad (18.5) \]

We have used a Maxwell’s relation:
\[ \frac{\partial(S, L)}{\partial(F, L)} = -1. \]

One more for a gas: How about the sign of
\[ \frac{\partial S}{\partial P} \bigg|_V \quad (18.6) \]

To increase \(P\) under constant \(V\), (usually) we have to raise the temperature, resulting in the increase of entropy, so the sign must be positive. The cleverest approach may be
\[ \frac{\partial S}{\partial P} \bigg|_V = \frac{\partial(S, V)}{\partial(P, V)} = \frac{\partial(S, V)}{\partial(T, -P)} \frac{\partial(T, -P)}{\partial(P, V)} = \frac{\partial(S, V)}{\partial(T, -P)} \bigg/ \frac{\partial V}{\partial T} \bigg|_P. \quad (18.7) \]

Therefore, \((18.6)\) and \((\partial V/\partial T)_P\) have the same sign, but to understand this statement we need the following inequality resulting from the system stability:
\[ \frac{\partial(S, V)}{\partial(T, -P)} > 0. \quad (18.8) \]

For a gas \((\partial V/\partial T)_P > 0\) without doubt, but this sign is not due to some reason of principle nature (in contradistinction to the inequality \((18.8)\)). For liquid water below 4°C under the atmospheric pressure, this derivative is indeed negative.

**18.2 Two kinds of inequalities, sacred and not**

As we will see in this section we encounter two different types of inequalities in thermodynamics; one class is due to some reason of thermodynamic principle, and the
other is only due to materialistic accident. Generally speaking, what is not forbidden by thermodynamics does happen in our world. According to our ordinary experiences (and also due to the microscopic picture of materials) \((\partial V/\partial T)_P > 0\) looks quite natural, but thermodynamics does not say anything about this sign. That is, even if it is negative, thermodynamics would not complain, and indeed what is not forbidden by thermodynamics does happen in this case.

### 18.3 Universal stability criteria

Clausius told us that if a spontaneous change occurs in an isolated system,

\[
\Delta S \geq 0. \tag{18.9}
\]

We use the standard trick to study a non-isolated system S as a small part of a huge isolated system (Fig. 18.1) whose intensive variables are kept constant, but their conjugate extensive variables may be exchanged freely between S and its surrounding reservoir.

![Figure 18.1: S is a part of a huge isolated system whose intensive parameters \(T_e, P_e, x_e, \text{etc.},\) are kept constant. This is virtually possible because the whole system is huge. Their conjugate extensive quantities \(S\) (or heat), \(V, X, \text{etc.},\) can be freely exchanged between the system S and the rest.](image)

If something spontaneous can happen, the total entropy must increase. In the system something irreversible might have happened, so we cannot compute \(\Delta S\) directly with the aid of imported quantities \(\Delta E, \Delta V, \text{etc.}\) However, for the reservoir, since we assume it is always in equilibrium, we can write its entropy change as

\[
\Delta S_{res} = -\frac{1}{T_e} \Delta E - \frac{P_e}{T_e} \Delta V + \frac{x_e}{T_e} \Delta X. \tag{18.10}
\]

Here, \(\Delta E, \text{etc.},\) are the quantities seen from the system S (‘+ signs’ for importing to S), so \(-\Delta E, -\Delta V, \text{etc.},\) are the imported quantities to the reservoir. That is why the signs in (18.10) are different from the usual Gibbs relation. Thus, the total entropy change is \(\Delta S + \Delta S_{res},\) and the Clausius’ inequality for the isolated system reads

\[
\Delta S - \frac{1}{T_e} \Delta E - \frac{P_e}{T_e} \Delta V + \frac{x_e}{T_e} \Delta X + \frac{\mu_e}{T_e} \Delta N \geq 0. \tag{18.11}
\]
Here, although we have not yet discussed the change of number $N$ of particles and its conjugate variable $\mu$ (chemical potential), since it is formally quite similar to other terms, for the later convenience, the last term is added, which will be discussed in Lecture 19.

If the equilibrium state is stable, then
\[
\Delta S - \frac{1}{T_e} \Delta E - \frac{P_e}{T_e} \Delta V + \frac{x_e}{T_e} \Delta X + \frac{\mu_e}{T_e} \Delta N < 0. \tag{18.12}
\]

Now, let us look at $\Delta S$ more closely. If the changes are very small, we can Taylor-expand $\Delta S$ into a power series of $\delta E = \Delta E$, $\delta V = \Delta V$, etc. (here $\Delta$ for the independent variables is replaced by $\delta$ to make it clear that all changes are small). We can separate the entropy change into the first order small quantity $\delta S$, the second order small quantity $\delta^2 S$, etc., as
\[
\Delta S = \delta S + \delta^2 S + \cdots. \tag{18.13}
\]
The first order term reads
\[
\delta S = \frac{1}{T_e} \delta E + \frac{P_e}{T_e} \delta V - \frac{x_e}{T_e} \delta X - \frac{\mu_e}{T_e} \delta N, \tag{18.14}
\]
because the derivatives are computed around the equilibrium state. Combining this expression, (18.12) and (18.13), we conclude that the stability condition of the equilibrium state is
\[
\delta^2 S < 0 \tag{18.15}
\]
irrespective of the constraints imposed on the system $S$ (that is, independent of whether some extensive quantities are allowed to be exchanged or not). Thus, this is the universal stability condition for the equilibrium state.

Notice that (18.15) was concluded for isolated systems before as the max entropy principle, but here it is about the general non-isolated system, so this does not imply max entropy; irrespective of $S$ being max or not, $\delta^2 S < 0$ is the stability condition.

### 18.4 Universal stability criterion in terms of internal energy

(18.12) may be rearranged as
\[
\Delta E > T_e \Delta S - P_e \Delta V + x_e \Delta X + \mu_e \Delta N. \tag{18.16}
\]
Now, restricting the variations to small ones, we can Taylor-expand $\Delta E$ just as we did for $\Delta S$. You should immediately realize that a very similar logic as above can give us another, but equivalent universal stability criterion
\[
\delta^2 E > 0. \tag{18.17}
\]
18.5 Le Chatelier’s principle

Let us study the consequences of the stability criterion (18.17): a general expression is

$$\sum_{i,j} \frac{\partial^2 E}{\partial X_i \partial X_j} \delta X_i \delta X_j > 0.$$  

This is a positive definite quadratic form, and we can express it as

$$\begin{pmatrix} \frac{\partial T}{\partial S} |_{V,N} & \frac{\partial T}{\partial V} |_{S,N} & \frac{\partial T}{\partial N} |_{S,V} \\ -\frac{\partial P}{\partial S} |_{V,N} & -\frac{\partial P}{\partial V} |_{S,N} & -\frac{\partial P}{\partial N} |_{S,V} \\ \frac{\partial u}{\partial S} |_{V,N} & \frac{\partial u}{\partial V} |_{S,N} & \frac{\partial u}{\partial N} |_{S,V} \end{pmatrix} \begin{pmatrix} \delta S \\ \delta V \\ \delta N \end{pmatrix} > 0.$$  

Let us assume $N$ is constant for simplicity. The sign of (18.19) must always be positive irrespective of the choice of $\delta S$ and $\delta V$ (unless both are zero). Therefore, all the diagonal terms must be positive:

$$0 < \left. \frac{\partial T}{\partial S} \right|_V = \frac{T}{C_V},$$  

and in terms of the adiabatic compressibility $\kappa_S = -(\partial V/\partial P)_S/V$

$$0 < -\left. \frac{\partial P}{\partial V} \right|_S = \frac{1}{V \kappa_S}.$$  

You must imagine what happens if these signs are flipped. The diagonal inequalities are called Le Chatelier’s principle.\(^{227}\) We can verbally state the consequence as follows:

In equilibrium changes occur in the direction to discourage further changes (to avoid run-away processes).

For example, if $\Delta S$ is injected (e.g., heat is injected) into the system, its temperature goes up, which usually discourages further injection of heat. In the case of compressibility, decrease of the volume of the system increases the pressure, resisting further squishing. Thus, no runaway phenomenon is realized in the world we live in.

18.6 Le Chatelier-Braun’s principle
Which is larger, the susceptibility under constant extensive quantity and that under constant intensive quantity? An example is: which is larger \( C_P \) or \( C_V \)? A general answer is called the Le Chatelier-Braun’s principle. Since

\[
\frac{\partial X}{\partial x} \bigg|_Y = \frac{\partial X}{\partial y} \bigg|_x + \frac{\partial X}{\partial y} \bigg|_x \frac{\partial y}{\partial x} \bigg|_Y
\]

(18.22)

\[
= \frac{\partial X}{\partial x} \bigg|_y + \frac{\partial X}{\partial y} \bigg|_x \frac{\partial (y, Y)}{\partial (x, Y)}
\]

(18.23)

\[
= \frac{\partial X}{\partial x} \bigg|_y + \frac{\partial X}{\partial y} \bigg|_x \frac{\partial (y, Y)}{\partial (x, x)} \frac{\partial (x, Y)}{\partial (y, x)}
\]

(18.24)

\[
= \frac{\partial X}{\partial x} \bigg|_y - \frac{\partial X}{\partial y} \bigg|_x \frac{\partial y}{\partial Y} \bigg|_x
\]

(18.25)

Here, a Maxwell’s relation has been used. This implies that

\[
\frac{\partial X}{\partial x} \bigg|_y > \frac{\partial X}{\partial x} \bigg|_Y.
\]

(18.26)

Therefore, the indirect change in extensive quantities (of \( Y \) in the above formula) occurs in the direction to reduce the effect of the perturbation (Le Chatelier-Braun’s principle).\(^{228}\) As already noted, a typical example is \( C_P \geq C_V \): larger specific heat implies that it is harder to warm up, that is, the system becomes harder to heat up if the volume change (i.e., the indirect change) is allowed.

18.7 2 \times 2 stability criterion
A necessary and sufficient condition for (18.19) is the positivity of all the principal

\[^{228}\text{Karl Ferdinand Braun (1850-1918): Z. physik. Chen., 1, 269 (1887), Ann. Physik, 33, 337 (1888) [the inventor of the cathode-ray tube, the discoverer of principle of semiconductor diode, shared the Nobel prize with Marconi for wireless technology]. The history of this principle can be found in J. de Heer, “The principle of le Chatelier and Braun,” J. Chem. Educ., 34, 375 (1957). The form stated here is due to Ehrenfest.\}^\]

\[^{[1888: Clausius died; van Gogh cuts off his left ear, The Night Cafe; Paul Gauguin Vision After the Sermon; Tchaikovsky, 5th Symphony; Mahler No 2 Resurrection (Jansons, Concertgebouw).]}^\]
minors\(^{229}\) of the matrix in (18.19). Therefore, in particular,
\[
\frac{\partial (T, -P)}{\partial (S, V)} > 0. \tag{18.27}
\]
Generally,
\[
\frac{\partial (X, Y)}{\partial (x, y)} > 0, \tag{18.28}
\]
where \((x, X)\) and \((y, Y)\) are conjugate pairs. This is perhaps the last formula you should remember when you use the Jacobian technique.

### 18.8 Importance of fluctuations

We already know that the mesoscopic world is dominated by fluctuations, which allow us to have a glimpse of the atomic world underlying the world we experience daily. We also know that the equilibrium state of a macroscopic systems is always tested by fluctuations; \(\delta\) in the stability criterion is actually spontaneously realized by thermal fluctuations.

Thus, there is no doubt about the importance of fluctuations qualitatively. How about quantitatively? We will see that the system response to perturbation is quantitatively related to fluctuations.

### 18.9 Fluctuation-response relation: preparation

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). We wish to study the susceptibility
\[
\chi = \left. \frac{\partial X}{\partial x} \right|_{T, \cdots}. \tag{18.29}
\]
Here, \(\cdots\) depends on the system we study. Since we wish to use \(T, x, \cdots\) as independent variables, we wish to have a system that can freely exchange their conjugate extensive quantities with its environment. Then, it is convenient to use the thermodynamic potential \(\tilde{G}\) defined by the following Legendre transformation:
\[
E \rightarrow \tilde{G} = E - TS - xX = A - xX. \tag{18.30}
\]

\(^{229}\) You sample the same row and column numbers (say, 1, 3, 7 and 8th columns and 1, 3, 7 and 8th rows from the original matrix and make a determinant \(\det(a_{ij})\), where \(i, j \in \{1, 3, 7, 8\}\)). Such determinants are called principal minors.
\( \tilde{G} \) is a *generalized Gibbs free energy*. Why is this convenient? Recall the original Gibbs free energy \( G = A + PV \) (do not forget that the conjugate variable of \( V \) is \( -P \)), for which \( T \) and \( P \) are independent variables, and we have \( dG = -SdT + VdP \).

If we wish to use statistical mechanics, entropy is more fundamental and convenient, so we need the corresponding Legendre transformation for entropy:

\[
S \rightarrow -\tilde{G}/T = S - E/T + xX/T = -A/T + xX/T.
\]  

(18.31)

How can we directly compute \( \tilde{G} \) statistical mechanically?

Compare the following relations:

\[
S/k_B = \log w(E, X) \leftrightarrow w(E, X)
\]

(18.32)

\[
S/k_B - E/k_B T = -A/k_B T = \log Z(T, X) \leftrightarrow Z(T, X) = \sum_E w(E, X) e^{-E/k_BT}.
\]

(18.33)

You must have already guessed:

\[
\frac{1}{k_B} (S - E/T + xX/T) = -\frac{\tilde{G}}{k_B T} = \log \tilde{Z}(T, x) \leftrightarrow \tilde{Z}(T, x) = \sum_E w(E, X) e^{(-E + xX)/k_BT}.
\]

(18.34)

Notice that

\[
\tilde{Z}(T, x) = \sum_x Z(T, X) e^{xX/k_BT} = \sum_x e^{-\beta(H-xX)},
\]

(18.35)

and

\[
\tilde{G}(T, x) = -k_B T \log \tilde{Z}(T, x).
\]

(18.36)

\( \tilde{Z} \) is called a *generalized canonical partition function*. These relations just look like the one we are very familiar with.

**18.10 Fluctuation-response relation**

The susceptibility \( \chi = (\partial X/\partial x)_{T,V} \) of the response \( X \) to the change of \( x \) reads

\[
\chi = \beta \frac{\partial^2 \log \tilde{Z}}{\partial (\beta x)^2}.
\]

(18.37)

Let us compute this with the aid of the expression of the generalized canonical partition function. First, we obtain

\[
X = \frac{\partial \log \tilde{Z}}{\partial \beta x} = \frac{1}{\tilde{Z}} \sum \dot{X} e^{-\beta H + \beta x \dot{X}}.
\]

(18.38)
Let’s differentiate this once more:

\[
\chi = \beta \frac{\partial X}{\partial \beta x} = -\frac{1}{Z^2} \left( \sum \hat{X} e^{-\beta H + \beta x \hat{X}} \right)^2 + \frac{1}{Z} \sum \hat{X}^2 e^{-\beta H + \beta x \hat{X}}.
\] (18.39)

That is,

\[
\chi = \beta \left( \langle \hat{X}^2 \rangle - \langle \hat{X} \rangle^2 \right) = \beta \langle \delta \hat{X}^2 \rangle \geq 0,
\] (18.40)

where \( \delta \hat{X} = \hat{X} - \langle \hat{X} \rangle \).

\[
\chi \left( = \beta \frac{\partial X}{\partial \beta x} \right) = \beta \langle \delta \hat{X}^2 \rangle
\] (18.41)

is called the fluctuation-response relation.

If we have many variables, what can we have? Probably, you guessed the following form:

\[
\chi_{ij} \left( = \beta \frac{\partial X_i}{\partial \beta x_j} \right) = \beta \langle \delta \hat{X}_i \delta \hat{X}_j \rangle.
\] (18.42)

This is correct.

### 18.11 Three key outcomes of fluctuation-response relation

We can make three important observations from the fluctuation-response relation (18.41):

(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, \( N \)). Therefore, \( \delta \hat{X} = O(\sqrt{N}) \).\(^{230}\)

(iii) \( \chi \) cannot be negative. This is the manifestation of the stability of the equilibrium state as we have already discussed.

Can you guess the general formula when there are many variables?

Thus, we have realized that studying fluctuation is quite important; it could be a non-invasive method to study the system response. The study of fluctuation is a mesoscopic scale study of the system, so it is the study of large deviation. Thus, if we know the large deviation function \( I \), we are done. Einstein gave \( I \) for equilibrium fluctuations. This we will discuss in the next lecture.

\(^{230}\)away from critical points. There, \( \chi \) can diverge, so nothing can be said from this argument.
Q18-1.
(1) What is the sign of
\[
\frac{\partial S}{\partial V} \bigg|_P
\]
for liquid water below 4 °C (under the atmospheric pressure). You may use the empirical fact \( (\partial V/\partial T)_P < 0 \).
(2) Using the experimental result you confirmed for a rubber band (i.e., \( (\partial T/\partial L)_S > 0 \)), find the sign of
\[
\frac{\partial L}{\partial S} \bigg|_F.
\]
Then, give an intuitive explanation of your result. You may use the fact \( (\partial X/\partial x)_Y > 0 \) for any conjugate pair \( (x, X) \) and for any variable \( Y \) (need not be extensive).
(3) There is an elastic body for which \( (\partial S/\partial L)_T > 0 \). Find the sign of
\[
\frac{\partial L}{\partial T} \bigg|_F,
\]
where \( F \) is the tensile force, and \( L \) is the length. You may use the fact \( (\partial X/\partial x)_Y > 0 \) for any conjugate pair \( (x, X) \) and for any variable \( Y \) (need not be extensive, even though it is in the upper case.).

Soln.
I strongly recommend you to use the Jacobian technique, since you do not need any insight. The following (1) explains step by step how this technique works.

(1) First, write the partial derivative in terms of a Jacobian and then split the denominator and the numerator:
\[
\frac{\partial S}{\partial V} \bigg|_P = \frac{\partial (S, P)}{\partial (V, P)} = \frac{\partial (S, P)}{\partial (T, P)} \frac{\partial (T, P)}{\partial (V, P)}.
\]

Now, you must look at what is given or what you wish: it is \( (\partial V/\partial T)_P < 0 \):
\[
\frac{\partial (V, P)}{\partial (T, P)} = 1 / \frac{\partial (T, P)}{\partial (V, P)}.
\]

Since (18.46) already contains \( (V, P) \), we should introduce \( (T, P) \):
\[
\frac{\partial (S, P)}{\partial (V, P)} \frac{\partial (T, P)}{\partial (V, P)} = \frac{\partial (S, P)}{\partial (T, P)} \frac{\partial (T, P)}{\partial (V, P)} = \frac{C_p}{T} \frac{\partial V}{\partial T} \bigg|_P < 0,
\]
since the specific heat is positive, or the diagonal terms of the Hessian of $E$ is positive.

(2)

$$\frac{\partial L}{\partial S} \bigg|_F = \frac{\partial (L, F)}{\partial (S, F)} = \frac{\partial (L, F)}{\partial (T, S)} \frac{\partial (T, S)}{\partial (S, F)} \tag{18.49}$$

$$= \frac{\partial (T, S)}{\partial (L, S)} \frac{\partial (L, S)}{\partial (S, F)} = - \frac{\partial T}{\partial L} \bigg|_S \frac{\partial L}{\partial F} \bigg|_S < 0. \tag{18.50}$$

Here, a Maxwell’s relation

$$\frac{\partial (L, F)}{\partial (T, S)} = 1 \tag{18.51}$$

has been used.

(3)

$$\frac{\partial L}{\partial T} \bigg|_F = \frac{\partial (L, F)}{\partial (T, F)} = \frac{\partial (L, F)}{\partial (T, S)} \frac{\partial (T, S)}{\partial (T, F)} \tag{18.52}$$

$$= \frac{\partial (T, S)}{\partial (T, L)} \frac{\partial (T, L)}{\partial (T, F)} > 0. \tag{18.53}$$

Q18-2.

(1) Consider a rubber band whose Gibbs relation is $dE = TdS + FdL$ (as usual). Which is larger, the constant length specific heat $C_L$ or the constant force specific heat $C_F$? You may be able to guess this, but I ask you to prove this using the stability criterion.

(2) For a fluid which is larger, the isothermal compressibility $\kappa_T$ or the adiabatic compressibility $\kappa_S$? You may be able to guess this, but I ask you to prove this using the stability criterion.

Soln.

(1) Le Chatelier-Braun tells $C_L \leq C_F$. Let us demonstrate this, using

$$dS = \frac{\partial S}{\partial T} \bigg|_F dT + \frac{\partial S}{\partial F} \bigg|_T dF. \tag{18.54}$$

This implies (we wish to have the red factor squared)

$$\frac{\partial S}{\partial T} \bigg|_L = \frac{\partial S}{\partial T} \bigg|_F + \frac{\partial S}{\partial F} \bigg|_T \frac{\partial F}{\partial T} \bigg|_L \tag{18.55}$$

$$= \frac{\partial S}{\partial T} \bigg|_F + \frac{\partial S}{\partial F} \bigg|_T \frac{\partial (F, L)}{\partial T, L} \tag{18.56}$$
\[
\frac{\partial S}{\partial T}
\left|_F + \frac{\partial S}{\partial F}ight|_T \frac{\partial (F, L)}{\partial (S, T)} \frac{\partial (S, T)}{\partial (F, T)} \frac{\partial (S, T)}{\partial (F, L)}
\]

\[= \frac{\partial S}{\partial T}
\left|_F + \frac{\partial S}{\partial F}ight|_T \frac{\partial (F, T)}{\partial (T, L)}^2 \] (18.57)

where we used a Maxwell’s relation

\[\frac{\partial (F, L)}{\partial (S, T)} = 1.\] (18.59)

Now,

\[\frac{\partial S}{\partial T}
\left|_L = \frac{\partial S}{\partial T}
\left|_F - \frac{\partial S}{\partial F}
\left|_T \frac{\partial F}{\partial L} \right|_T.\] (18.60)

Thanks to the stability the second term is negative, so

\[\frac{\partial S}{\partial T}
\left|_L \leq \frac{\partial S}{\partial T}
\left|_F.\] (18.61)

Therefore,

\[C_L \leq C_F.\] (18.62)

You must think whether this is natural or not. Take a polymer chain or a rubber band. Under constant length, increasing the temperature increases the force, which opposes chain conformational entropy increase. In contrast, under constant force, more conformational varieties are available than the former case, so more heat is needed to increase the system temperature. Thus, \(C_L\) should be smaller than \(C_F\). A parallel argument should be possible for systems with energetic elasticity. Try.

Let us follow the above strategy to show \(C_V \leq C_P\) for practice sake.

\[\frac{\partial S}{\partial T}
\left|_V = \frac{\partial S}{\partial T}
\left|_P + \frac{\partial S}{\partial P}
\left|_T \frac{\partial P}{\partial V} \right|_V \] (18.63)

\[= \frac{\partial S}{\partial T}
\left|_P + \frac{\partial S}{\partial P}
\left|_T \frac{\partial (P, V)}{\partial (T, V)} \] (18.64)

\[= \frac{\partial S}{\partial T}
\left|_P + \frac{\partial S}{\partial P}
\left|_T \frac{\partial (P, V) \partial (S, T)}{\partial (P, T) \partial (T, V)} \] (18.65)

\[= \frac{\partial S}{\partial T}
\left|_P - \frac{\partial S}{\partial P}
\left|_T \frac{\partial (P, T)}{\partial (T, V)} \] (18.66)

\[= \frac{\partial S}{\partial T}
\left|_P + \frac{\partial S}{\partial P}
\left|_T \frac{\partial P}{\partial V} \right|_T.\] (18.67)
where we used a Maxwell’s relation (note \(-P\) is the conjugate of \(V\))

\[
\frac{\partial(P, V)}{\partial(S, T)} = -1. \tag{18.68}
\]

Thus,

\[
C_V \leq C_P, \tag{18.69}
\]

because \(\partial P/\partial V < 0\).

(2) Let us start with

\[
dV = \left. \frac{\partial V}{\partial P} \right|_T dP + \left. \frac{\partial V}{\partial T} \right|_P dT, \tag{18.70}
\]

Thus,

\[
\left. \frac{\partial V}{\partial P} \right|_S = \left. \frac{\partial V}{\partial P} \right|_T + \left. \frac{\partial V}{\partial T} \right|_P \left. \frac{\partial T}{\partial P} \right|_S \tag{18.71}
\]

\[
= \left. \frac{\partial V}{\partial P} \right|_T + \left. \frac{\partial V}{\partial T} \right|_P \frac{\partial(T, V)}{\partial(V, P)} \frac{\partial(T, P)}{\partial(P, S)} \tag{18.72}
\]

\[
= \left. \frac{\partial V}{\partial P} \right|_T - \left. \frac{\partial V}{\partial T} \right|_P \left. \frac{\partial(T, P)}{\partial(P, S)} \right|^2. \tag{18.73}
\]

where a Maxwell relation

\[
\frac{\partial(T, S)}{\partial(V, P)} = -1 \tag{18.74}
\]

is used.

\[
\left. \frac{\partial V}{\partial P} \right|_S = \left. \frac{\partial V}{\partial P} \right|_T + \left. \frac{\partial V}{\partial T} \right|_P \left. \frac{\partial T}{\partial S} \right|_P. \tag{18.75}
\]

Hence,

\[
\left. \frac{\partial V}{\partial P} \right|_S \geq \left. \frac{\partial V}{\partial P} \right|_T. \tag{18.76}
\]

Notice these derivatives are negative, and compressibilities are defined with negative signs, so

\[
\kappa_S \leq \kappa_T. \tag{18.77}
\]

Is this natural? We know that under adiabatic conditions if we compress an ideal gas \(PV^\gamma = \text{constant} (\gamma > 1)\) or \(V \propto 1/P^{1/\gamma}\). Under isothermal conditions \(PV\) is constant, or \(V \propto 1/P\). Therefore, \(V\) is smaller under isothermal compression. That is, it is easier to squish under const \(T\) than const \(S\). More intuitively, adiabatic compression usually increases the gas temperature, so isothermal compression should be easier.
Q18-3. We wish to study the fluctuation of the total energy $E$ of a closed system (a system without any material exchange) under constant volume $V$ and temperature $T$ (i.e., thermostatted). To what heat capacity of the system (say, under constant pressure or constant volume) does $\langle \delta E^2 \rangle$ directly related? Notice that microscopically $E$ is just the system Hamiltonian $\mathcal{H}$, so we are interested in $\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2$. The most straightforward way to answer this question is to compute this variance.

**Soln.**

$Z = \sum e^{-\beta \mathcal{H}}$, so

$$\langle \mathcal{H} \rangle = \frac{\partial \log Z}{\partial (-\beta)} | _V = \frac{1}{Z} \sum \mathcal{H} e^{-\beta \mathcal{H}}.$$  

Therefore,

$$\frac{\partial \langle \mathcal{H} \rangle}{\partial (-\beta)} | _V = -\frac{1}{Z^2} \left( \sum \mathcal{H} e^{-\beta \mathcal{H}} \right)^2 + \frac{1}{Z} \sum \mathcal{H}^2 e^{-\beta \mathcal{H}} = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2.$$  

We know

$$\frac{\partial \langle \mathcal{H} \rangle}{\partial (-\beta)} | _V = k_B T^2 \frac{\partial E}{\partial T} | _V = k_B T^2 C_V.$$
19 Thermodynamic approach to fluctuations

Summary
* Einstein gave a universal probability distribution for mesoscopic scale fluctuations around equilibrium states.
* Thermodynamic stability and equilibrium work needed to create fluctuations are closely related to this universal distribution.

Key words
fluctuation-response relation, large deviation rate function, fundamental formula for thermodynamic fluctuations (Einstein’s theory), multivariate Gaussian distribution

What you should be able to do
* Explain why Einstein’s theory is always correct.
* Computation of second moments of fluctuations, or how to use Einstein’s theory
* How to use multivariate Gaussian distribution; you must be able to compute its normalization constant.

19.1 Mesoscopic fluctuations: introduction to Einstein’s theory
The story up to this point is about a whole finite system. You may regard it as a mesoscopic subsystem of a larger system that is of our daily scale. If this is what we want, however, there is a reason that we may avoid the detailed statistical-mechanical setup just we have discussed. Recall the universal stability criterion for an equilibrium state \( \delta^2 S < 0 \). This is independent of the environmental constraints, and it is a condition preventing fluctuations from going wild. Then, there should be a theory that can determine the distribution of fluctuations almost independent of the environmental constraints imposed on the system. As we will see soon, Einstein just constructed such a theoretical framework.

The study of fluctuation is a mesoscopic scale study of the system, so it is the study of large deviation. Since we study a small volume \( V \) in the system, the following type of large deviation must be natural:

\[
P \left( \frac{1}{V} X(V) \sim y \right) \approx e^{-V I(y)}, \tag{19.1}
\]

where \( I \) is the large deviation function (or rate function).\(^2\) If we know \( I \), basically

\(^2\) If \( X \) is extensive, \( X(V) \) is the total amount in \( V \) (i.e., \( X(V)/V \) is its density). If \( X \) is intensive,
we know everything we wish to know about fluctuations.

In practice, 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. That is, we need the quadratic approximation to $I$, which was provided by Einstein.

19.2 Einstein’s fundamental formula for small fluctuations

Einstein in 1910 studied the deviation of thermodynamic observables in a small domain of a system from their equilibrium values in order to understand critical fluctuations, which we will discuss towards the end of this course.

To obtain the probability of fluctuations, he inverted the Boltzmann principle as

$$w(\{X\}) = e^{S(\{X\})/k_B},$$

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 (19.2). Since we know the statistical weights, we can compute the probability of observing $\{X\}$ as

$$P(\{X\}) = \frac{w(\{X\})}{\sum_{\{X\}} w(\{X\})}.$$  (19.3)

The denominator may be replaced with the largest term in the summands (recall 14.4), so we may rewrite the formula as

$$P(\{X\}) \simeq \frac{w(\{X\})}{w(\{X_{eq}\})} = e^{[S(\{X\}) - S(\{X_{eq}\})]/k_B} = e^{-|\Delta S|/k_B},$$

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, then $X(V)/V$ should be interpreted as the average value in the volume $V$.

232[1910: Russel and Whitehead started to publish Principia Mathematica (∼1913); Stravinsky, Firebird premiered; Rilke, Die Aufzeichnungen des Malte Laurids Brigg; the Mexican revolution began, Tolstoy, Cannizzaro, Nightingale died.]

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 this reads

$$P(\{\delta X\}) \propto e^{-|\delta^2 S|/k_B}.$$  \hspace{1cm} (19.5)

Einstein proposed this as the fundamental formula for small fluctuations in a small portion of any equilibrium system.

The above derivation of (19.5) assumed that a mesoscopic portion of the system is isolated. However, as noted before, the second order deviation of any thermodynamic potential around its equilibrium point is given by $-T\delta^2 S$, so (19.5) is valid under any condition; Einstein is always correct.\(^{234}\)

### 19.3 Practical form of fluctuation probability

To study the fluctuation we need the second order variation $\delta^2 S$ of the system entropy. 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)

$$\delta S = \frac{1}{T}(\delta E + P\delta V - \mu\delta N - x\delta X)$$  \hspace{1cm} (19.6)

as follows (this is the second order term of the Taylor expansion, so do not forget

\(^{234}\)If you do not trust Einstein, use the ‘generalized canonical partition function’ (appropriate for the constraints imposed on the small portion) as

$$\tilde{Z}(X) = e^{S/k_B - xX/k_BT},$$

and compute the probability for the fluctuation $\delta X$: $P(\delta X)$.  

274
the overall factor $1/2$):  

$$
\delta^2 S = \frac{1}{2} \left[ \delta \left( \frac{1}{T} \right) (\delta E + P \delta V - \mu \delta N - x \delta X) + \frac{1}{T} (\delta P \delta V - \delta \mu \delta N - \delta x \delta X) \right],
$$

(19.7)

$$
= -\frac{\delta T}{2T^2} T \delta S + \frac{1}{2T} (\delta P \delta V - \delta \mu \delta N - \delta x \delta X).
$$

(19.8)

Thus, we have arrived at the following useful expression worth remembering (actually, almost nothing to remember anew; cf. $-\delta E = -T \delta S + P \delta V - \mu \delta N - x \delta X$):  

$$
\delta^2 S = \frac{-\delta T \delta S + \delta P \delta V - \delta \mu \delta N - \delta x \delta X}{2T}.
$$

(19.9)

Don’t forget 2 downstairs.

Consequently, the probability density of fluctuation can have the following form, which is the starting point of practical calculation of fluctuations (second moments):

$$
P(\text{fluctuation}) \propto \exp \left\{ -\frac{1}{2k_B T} (\delta T \delta S - \delta P \delta V + \delta \mu \delta N + \delta x \delta X) \right\}.
$$

(19.10)

---

235If 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., $\delta^2 x = \delta^2 y = 0$, everybody must know this):

$$
\delta [X \delta x + Y \delta y] = \delta X \delta x + X \delta^2 x + \delta Y \delta y + Y \delta^2 y = \delta X \delta x + \delta Y \delta y.
$$

236As has already been stated when we discussed the general stability criterion, $\delta S$ need not be zero. The derivation of this formula by Einstein was indeed a feat.
19.4 Fluctuation and reversible work needed to create it

A similar calculation gives

$$\delta^2 E = \frac{1}{2}(\delta T \delta S + \delta P \delta V - \delta \mu \delta N - \delta x \delta X).$$

(19.11)

$\delta^2 E$ can be understood as the (free) energy we must supply as the reversible work, if we wish to create the fluctuation. Therefore, we may rewrite (19.10) as

$$P(\text{fluctuation}) \propto e^{-\beta W_f},$$

(19.12)

where $W_f$ is the reversible work required to create the fluctuation. This is a practically very useful formula.

19.5 How to use the practical formula

To use the practical formula (19.10) we must choose the independent variables.

What variables should we choose 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). Such a system requires $n$ conjugate pairs $(S, T), (V, -P), (x_i, X_i)$, etc. Choosing a statistical ensembles corresponds to choosing one variable from each pair. We know, however, any ensemble may be used to study the second moments of thermodynamic fluctuations. Thus, we may choose $n$ independent variables arbitrarily 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.

After choosing the independent variables, the formula in the round parentheses of (19.10) becomes a quadratic form in independent variations (of our choice, say, $\{\delta T, \delta V, \delta x\}$).

For an example, let us calculate the temperature fluctuation. We must first choose independent variables (variations). $\delta T$ must be a convenient choice. We need one
more independent \( \delta V \) as the other variation. We could choose \( \delta P \), but we wish to exploit the following general fact:

\[
\langle \delta X_i \delta x_j \rangle = k_B T \delta_{ij},
\]

where \( X_i \) is an extensive variable, and \( x_j \) an intensive variable; we understand that \((X_i, x_i)\) is a conjugate pair.\(^{237}\)

Thus, for \( \delta T \delta V \) is a convenient partner:

\[
\frac{1}{k_B} \delta^2 S = -\frac{1}{2k_B T} \delta S \delta T + \cdots = -\frac{1}{2k_B T} \left. \frac{\partial S}{\partial T} \right|_V \delta T^2 + \cdots = -\frac{C_V}{2k_B T^2} \delta T^2 + \cdots. \tag{19.15}
\]

Therefore, we can easily conclude

\[
\langle \delta T^2 \rangle = k_B T^2 / C_V. \tag{19.16}
\]

This is the fluctuation of the average temperature in a small volume whose constant volume heat capacity is \( C_V \). If the observation volume is reduced, then \( C_V \) is reduced as well (recall that \( C_V \) is an extensive quantity), so the fluctuation increases. Very natural.

One more example: the pressure fluctuation. We should choose \( P \) and \( S \) as independent variables, since we know \( \langle \delta P \delta S \rangle = 0 \):

\[
\frac{1}{2k_B T} \delta P \delta V = \frac{1}{2k_B T} \left. \frac{\partial V}{\partial P} \right|_S \delta P^2 + \cdots = -\frac{V}{2k_B T} \kappa_S \delta P^2 + \cdots, \tag{19.17}
\]

where \( \kappa_S \) is the adiabatic compressibility

\[
\kappa_S = -\left. \frac{1}{V} \frac{\partial V}{\partial P} \right|_S. \tag{19.18}
\]

Therefore,\(^{238}\)

\[
\langle \delta P^2 \rangle = k_B T / V \kappa_S. \tag{19.19}
\]

\(^{237}\)\(^{237}\) This can be demonstrated with the aid of the fluctuation-response relation (18.42). Let us choose first \( X_1, \ldots, X_n \) as our independent variables. Using the chain rule, we obtain

\[
\langle \delta X_i \delta x_j \rangle = \left\langle \delta X_i \sum_k \frac{\partial x_j}{\partial X_k} \delta X_k \right\rangle = k_B T \sum_k \frac{\partial x_j}{\partial X_k} \frac{\partial X_k}{\partial x_i} = k_B T \frac{\partial x_j}{\partial x_i} = k_B T \delta_{ij}. \tag{19.14}
\]

In the above calculation, in the last partial derivative, independent variables are switched to \( x_1, \ldots, x_n \).

\(^{238}\)\(^{238}\) The fluctuations of the quantities that may be interpreted as the expectation values of microscopic-mechanically expressible quantities (e.g., internal energy, volume, pressure) tend to zero in the \( T \to 0 \) limit as we see here. However, the fluctuations of entropy and the quantities obtained by differentiating it do not satisfy the above property.
Such small tricks to save our energy are useful, but as is clear, generally speaking, we must know how to handle multivariate Gaussian distribution.

### 19.6 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, the distribution is Gaussian:

$$f(x) = \frac{1}{\sqrt{\det(2\pi V)}} \exp \left(-\frac{1}{2}(x - m)^T V^{-1} (x - m)\right),$$  \hspace{1cm} (19.20)

where \(\langle x \rangle = m\), the mean, and \(V\) is the covariance matrix defined as (do not forget that our vectors are column vectors)

$$V = \langle (x - m)(x - m)^T \rangle.$$  \hspace{1cm} (19.21)

The reader must not have any difficulty in demonstrating that (19.20) is correctly normalized. [Hint: choose eigendirections of \(V\) as the orthogonal coordinates.]

In particular, for the two variable case:

$$f(x, y) \propto \exp \left\{ -\frac{1}{2} \left( ax^2 + 2bxy + cy^2 \right) \right\},$$  \hspace{1cm} (19.22)

then

$$V = \Lambda^{-1} = \left(\begin{array}{cc} a & b \\ b & c \end{array}\right)^{-1} = \frac{1}{\det \Lambda} \left(\begin{array}{cc} c & -b \\ -b & a \end{array}\right).$$  \hspace{1cm} (19.23)

That is,

$$\langle x^2 \rangle = c/\det \Lambda, \quad \langle xy \rangle = -b/\det \Lambda, \quad \langle y^2 \rangle = a/\det \Lambda.$$  \hspace{1cm} (19.24)
Q19-1. There is a deformable macromolecule (due to conformational changes) whose end-to-end distance is $L$. The ambient temperature is $T = 300$ K.

(1) If the amplitude of the deformation is small, the system may be understood as a harmonic spring. What is the effective spring constant $k$ of this macromolecule, when the length variance is known to be $\langle \delta L^2 \rangle$? [Hint: $k$ is the inverse susceptibility of $L$ against $F$ under constant temperature.]

(2) Suppose the needed force $F$ to maintain the end-to-end length to be $L$ is

$$F = k(L - L_0) + \frac{1}{3} \alpha (L - L_0)^3,$$  \hspace{1cm} (19.25)

where $L_0$ and $\alpha$ are positive constants. What is the fluctuation of $L$ around $L = L_1$? Notice that in this way without measuring $F$ you could reconstruct the $F$-$L$ relation.

Soln.

(1) A fluctuation-response relation applied to our case around $F = 0$ is

$$k_B T \frac{\partial L}{\partial F} \bigg|_T = \langle \delta L^2 \rangle.$$  \hspace{1cm} (19.26)

Therefore,

$$\frac{k_B T}{k} = \langle \delta L^2 \rangle.$$  \hspace{1cm} (19.27)

You must have seen something like this for a polymer chain.

By the way, we discussed that the susceptibility or $\frac{\partial X}{\partial x}$ is extensive. In our example, $L$ is extensive, so $1/k$ must be extensive. It means that if the length of the spring is doubled, its Hooke’s constant is halved, as we know well.

(2) Using the fluctuation-response relation, we have when $L = L_1$ (by applying an appropriate force)

$$\langle \delta L^2 \rangle = k_B T \frac{\partial L}{\partial F} \bigg|_T = \frac{k_B T}{k + \alpha (L_1 - L_0)^2}.$$  \hspace{1cm} (19.28)

Notice that $\langle \delta L^2 \rangle$ here and that in (1) are different quantities, because the states around which we observe fluctuations are distinct. In (1) it is the unstretched state, but here, it is somehow stretched and its mean length is $L_1$. It is much easier to measure the position than the force, so experimentally, the derivative of $F$ wrt $L$ may be obtained by measuring only the positions under unspecified appropriate forces.

Q19-2. [Easy questions about fluctuations]

(1) Consider a small portion of a big system in equilibrium. Let us assume that
the small portion contains a constant number of particles. Find the fluctuation of
entropy \( \langle \delta S^2 \rangle \) in terms of an appropriate heat capacity of the small portion.

(2) Take a (constant volume) small portion of a solution of some substance in a
solvent. Let \( c \) be the concentration of the solute. Show that \( \langle \delta c \delta T \rangle = 0 \), that is,
concentration fluctuation and temperature fluctuation are uncorrelated in equilib-
rium.

**Soln**

(1) We use (needless to say, we use Einstein’s thermodynamic fluctuation theory)

\[-\frac{1}{2k_B T} \delta T \delta S + \cdots.\]

Choose \( S \) and \( P \) as independent variables.\(^{239}\) Then we can rewrite this as

\[-\frac{1}{2k_B T} \left. \frac{\partial T}{\partial S} \right|_P \delta S^2 + \cdots = -\frac{1}{2k_B C_P} \delta S^2 + \cdots.\]

Therefore, \( \langle \delta S^2 \rangle = k_B C_P \), where \( C_P \) is the constant pressure specific heat of the
small portion.

(2) This is obvious from \( \langle \delta N \delta T \rangle = 0 \).

\(^{239}\)You need not, but since \( \langle \delta S \delta P \rangle = 0 \), computation is (drastically) simplified.
20 Chemical potential

Summary
* To discuss open systems we must extend the first law of thermodynamics by including the mass action \( dZ' = \sum \mu dN \).
* When exchange of a chemical is allowed between various parts of the system, the equilibrium condition is the identity of its chemical potentials among the parts.
* The chemical potential generally has the form \( \mu = \mu^\ominus + k_B T \log a \), where activity \( a \) is related to the concentration of the chemical.
* Algebraic expression of chemical reactions gives the reaction equilibrium condition \( \sum \nu \mu = 0 \), which leads to the concept of equilibrium constant.

Key words
mass action, chemical potential, Gibbs-Duhem relation, phase equilibrium, Clapeyron-Clausius equation, osmotic pressure, van’t Hoff’s law, Raoult’s law, colligative properties, chemical reaction, signed stoichiometric coefficients, law of mass action, equilibrium constant, van’t Hoff’s equation

What you should be able to do
* Be able to explain what the chemical potential is, and understand various equilibrium conditions in terms of chemical potentials.
* Be able to understand the shifting direction of the reaction when \( T \) or \( P \) are altered.

20.1 Open systems
So far we have discussed isolated systems, thermally isolated systems, and systems whose work coordinates are buffered (e.g., thermostatted systems). These are not allowed to exchange substances between the systems under consideration and their surrounding world. They are (materially) closed systems. Today, we will discuss open systems for which exchange of their component chemicals between themselves and their environments is allowed. We must extend the first law in the following form
\[
\Delta E = Q + W + Z, \tag{20.1}
\]
where \( Z \) is called the mass action, which describes the energetic contribution of materials exchange.
20.2 Mass action and chemical potential

To begin with, for simplicity, we assume only one chemical may be exchanged. We need one variable $N$ to specify its amount in the system (in moles or in numbers\textsuperscript{240}). We prepare a semipermeable rigid membrane that allows the passage of this chemical. Setting up a device as illustrated in Fig. 20.1 and injecting this chemical (only) into the system (or sucking up from the system), we can measure the necessary work (necessary mass action) $d'Z$ to inject $dN$ molecules into the system. This is written as

$$d'Z = \mu dN. \quad (20.2)$$

![Semipermeable membrane and system](image)

Figure 20.1: How to measure a chemical potential; however, this is only a very schematic figure.

Since ordinary work coordinates and entropy must be kept constant, it is practically almost impossible to use this approach. Practically, we use (20.14) under $T$, $P$ (and other intensive conjugate variables to work coordinates) constant.

Here, it is explicitly noted that $Z$ is not a state function, and $\mu$ is called the chemical potential of this chemical.

If we have more than one chemicals that we can change independently,\textsuperscript{241} we can appropriately generalize the above setup and write

$$d'Z = \sum_i \mu_i dN_i, \quad (20.3)$$

where $N_i$ denotes the amount of the $i$th chemical and $\mu_i$ the corresponding chemical potential.

Intuitively, the chemical potential of a chemical of a system is the measure of the

\textsuperscript{240}But in these lecture notes, $N$ always implies the number of particles, and we will not use moles unless clearly stated. Thermodynamics only handles the situation that the atomic nature of the material is not discernible, but here we adopt an eclectic attitude.

\textsuperscript{241}What is independently changed and what not may not be a simple question, but usually, common-sense tells us the right answer. For example, if you want to inject water into the system, inevitably, you inject OH\textsuperscript{-}, H\textsubscript{3}O\textsuperscript{+}, etc. as well, but they are ‘slaved’ to the total amount of water when $T$ and $P$ are fixed (chemical equilibrium). Therefore, we may conclude that only 1 component matters, and this answer is in agreement with our common sense conclusion.
system’s capability to export the chemical.

20.3 Gibbs relation
Including the mass action, the full form of the Gibbs relation reads

\[ dE = TdS - PdV + \mu dN + xdX + \cdots. \] (20.4)

or

\[ dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN - \frac{x}{T}dX + \cdots. \] (20.5)

Needless to say, if you have several independently modifiable chemicals, we must replace \( \mu dN \) with a sum over these chemicals \( \sum_i \mu_i dN_i \). Be careful about the signs.

Other thermodynamic potentials read

\[ dA = -SdT - PdV + \mu dN + xdX + \cdots, \] (20.6)

\[ dG = -SdT + VdP + \mu dN + xdX + \cdots. \] (20.7)

20.4 Gibbs-Duhem relation
If the system size is increased by a fraction \( \delta \lambda \) by joining a small fraction of the identical

\[ \begin{array}{c|c|c}
1 & \delta \lambda & 1 + \delta \lambda \\
\end{array} \]

Figure 20.2: A fraction \( \delta \lambda \) of the original system is added. Then, the increase of the extensive quantity \( X \) is \( \delta X = X \delta \lambda \) (i.e., \( X \rightarrow X + X \delta \lambda \)).

system in the identical equilibrium state (Fig. 20.2), all the extensive quantities are multiplied by \( 1 + \delta \lambda \). However, all the conjugate intensive quantities are intact. Therefore, (20.4) applied to this situation reads

\[ E \delta \lambda = (TS - PV + \mu N + xX + \cdots) \delta \lambda, \] (20.8)

or

\[ E = TS - PV + \mu N + xX + \cdots. \] (20.9)
This implies
\[ dE = (TdS - PdV + \mu dN + x dX + \cdots) + (SdT - VdP + Nd\mu + Xdx + \cdots), \quad (20.10) \]
but (20.4) is true, so we must conclude that
\[ SdT - VdP + Nd\mu + Xdx + \cdots = 0. \quad (20.11) \]
This relation is called the Gibbs-Duhem relation. This tells us how the chemical potential changes as a function of \( T \), \( P \), etc., if there is only one chemical species:
\[ d\mu = -\frac{S}{N}dT + \frac{V}{N}dP - \frac{X}{N}dx + \cdots. \quad (20.12) \]

### 20.5 Gibbs free energy in terms of chemical potentials
If we combine (20.9) with the definition of the Gibbs free energy, we obtain
\[ G = \mu N + x X + \cdots. \quad (20.13) \]
Usually, there are many chemicals and no other work coordinates than \( V \), so this becomes
\[ G = \sum_i \mu_i N_i. \quad (20.14) \]
If there is only one chemical, notice that the Gibbs relation (20.7) is just (20.12).

(20.13) can be more directly obtained by the same approach as is used to derive (20.9). We know
\[ dG = -SdT + VdP + \mu dN + x dx. \quad (20.15) \]
By the grafting process in Fig. 20.2, \( G \rightarrow G + G\delta \lambda \), since \( G \) is extensive, but since \( T \) and \( P \) are intact (intensive!). Therefore, (20.15) reads
\[ G\delta \lambda = \mu N\delta \lambda + x X\delta \lambda, \quad (20.16) \]
which is just (20.13).

### 20.6 Equilibrium condition
Suppose two systems I and II are joined to exchange heat, volume and chemicals while the whole system is in isolation (or thermally isolated and no work nor mass action is
provided from outside). Then, the equilibrium condition must be the maximization of the total entropy. Let $S_X$ be the entropy of system X. Then, $S = S_I + S_{II}$. If we assume thermal and pressure equilibration have already been attained, the remaining equilibrium condition is

$$\delta S = \sum_i \left( \frac{\mu_I}{T} \delta N_I + \frac{\mu_{II}}{T} \delta N_{II} \right) = 0. \quad (20.17)$$

Assuming that there is no chemical reaction changing $N_i$’s, we must conclude $\delta N_I + \delta N_{II} = 0$. Therefore, the equilibrium condition is

$$\mu_I = \mu_{II} \quad (20.18)$$

for each $i$ that can be exchanged between the two subsystems.

### 20.7 Phase equilibria

We will discuss phase transitions\(^{242}\) in detail toward the end of this lecture, but let us discuss some elementary and important aspects of phase co-existence associated with first order phase transitions (= discontinuous phase transitions).\(^{243}\)

If two phases coexist (just as ice floating on liquid water), we may regard different phases as different compartments I and II in contact with each other through the interphase (phase boundary). Let us study the condition for the equilibrium of these two phases (phase equilibria of these two phases) of a pure substance under constant $T$ and $P$. The equilibrium condition is the minimum of the Gibbs free energy of the whole system. Since there is only one chemical component,

$$0 = \delta G = \mu_I \delta N_I + \mu_{II} \delta N_{II}, \quad (20.19)$$

Therefore, if the system is materially closed, $\delta N = \delta N_I + \delta N_{II} = 0$, so we must conclude

$$\mu_I = \mu_{II}. \quad (20.20)$$

This is the phase equilibrium condition, which may be rewritten as

$$\Delta \mu = 0, \quad (20.21)$$

\(^{242}\)What is the phase transition? We will discuss the topic in depth later. Here, you have only to consider familiar examples such as melting of ice or boiling of water.

\(^{243}\)There are two major classes of phase transitions, continuous and discontinuous. In the former, there is no jump in any extensive quantities across the transition, but something strange can happen (say, their derivatives = susceptibilities diverge). In contrast, for discontinuous phase transitions at least one extensive quantity changes discontinuously at the phase transition point. For example, when ice melts, the density changes.
where $\Delta$ implies the change due to phase transition.

### 20.8 Clapeyron-Clausius equation

It is often interesting to know what happens, e.g., to the boiling point if the pressure is reduced (cf. vacuum distillation). To this end we must understand how the chemical potential changes. By the same logic used to derive the Gibbs-Duhem relation, we arrive at

$$d\mu = v dP - s dT,$$

where $v = V/N$ and $s = S/N$. These densities depend on the phases, but $T$ and $P$ are common to the coexisting phases, so (20.21) reads

$$\Delta v dP = \Delta s dT,$$

where $dT$ and $dP$ are changes along the phase coexistence line (the arrow in Fig. 20.3), and $\Delta v = v_1 - v_\Pi$ and $\Delta s = s_1 - s_\Pi$.

![Figure 20.3: What happens to the phase coexistence temperature if the phase coexistence pressure is changed by $dP$? Here, the red curve describes the phase transition line between phase I and phase II.](image)

### Figure 20.3

(20.23) implies

$$\left.\frac{dP}{dT}\right|_{\text{coexist}} = \frac{\Delta s}{\Delta v} = \frac{s_1 - s_\Pi}{v_1 - v_\Pi},$$

which is called the **Clapeyron-Clausius equation**. Here, you may also interpret $v$ as the molar volume (volume/one mole) and $s$ as the molar entropy. Since the entropy change $\Delta s$ is related to the latent heat $\Delta h_i = h_1 - h_\Pi$ as $\Delta s = \Delta h/T$ with the phase transition temperature $T$, we can also write

$$\left.\frac{dP}{dT}\right|_{\text{coexist}} = \frac{\Delta h}{T \Delta v}.$$
20.9 Chemical potential of ideal gas

Thus, we have learned that chemical potentials are fundamentally important. What can thermodynamics say about the chemical potential? From (20.22), if the temperature is fixed

\[ d\mu = \frac{V}{N} dP. \]  

(20.26)

Therefore, if we know the equation of state, we can say something about the chemical potential. For example, for an ideal gas we have

\[ d\mu = \frac{k_B T}{P} dP. \]  

(20.27)

That is,

\[ \mu(T, P) = \mu(T, P^\ominus) + k_B T \log(P/P^\ominus). \]  

(20.28)

Here \( \ominus \) denotes some standard state. In practice, we often write

\[ \mu(T, P) = \mu^\ominus(T) + k_B T \log P, \]  

(20.29)

where \( \mu^\ominus(T) \) is called the standard chemical potential.

Let us compute the chemical potential of the ideal gas statistical mechanically. We have computed the Helmholtz free energy, so we may obtain

\[ \mu = \frac{\partial A}{\partial N} \bigg|_{T,V}. \]  

(20.30)

We know

\[ Z = \frac{1}{N!} \left[ \left( \frac{2\pi m k_B T}{\hbar^2} \right)^{3/2} V \right]^N = \frac{1}{N!} \left[ \frac{\sqrt{2\pi}}{\lambda_T} L \right]^N = \left[ \frac{e n_Q V}{N} \right]^N, \]  

(20.31)

where we have used Stirling’s formula and \( n_Q = \left( \sqrt{2\pi} / \lambda_T \right)^3 \) with \( \lambda_T = (\hbar^2 / m k_B T)^{1/2} \), the thermal de Broglie wavelength.

We obtain

\[ A = N k_B T \log(N/V n_Q) - N k_B T, \]  

(20.32)

so

\[ \mu = k_B T \log \frac{N}{V n_Q} = k_B T \log \frac{n}{n_Q} = k_B T \log \frac{P}{k_B T n_Q}, \]  

(20.33)
where \( n \) is the number density. This indeed has the form (20.29).

The chemical potential of the \( i \)th component of the ideal gas mixture can be obtained with the aid of Dalton’s law of partial pressures:

\[
\mu_i(T, P) = \mu_i^\circ(T) + RT \log P_i,
\]

(20.34)

where \( P_i \) is the partial pressure of the \( i \)th gas.

### 20.10 Chemical potential of ideal solutions

Let us model the solvent/solution as a lattice model: each lattice point of the system is occupied either by a solvent molecule or by a solute molecule (Fig. 20.4).

![Figure 20.4: Lattice solution: \( N \) solvent molecules (yellow) and \( n \) solute molecules (blue).](image)

We ignore the interactions among molecules (other than the volume exclusion due to occupying the lattice points); we call such a solution an ideal solution. \( N \) solvent molecules are mixed with \( n \) solute molecules to make a solution containing an \( x = n/(N + n) \) mole fraction of solute. Let us discuss a dilute solution \( 0 < x \ll 1 \).

Let the chemical potential of a pure solvent be \( \mu_0(T, P) \) and that of a pure solute \( \mu_s(T, P) \). Then, the initial Gibbs free energy (before mixing) is \( G = N\mu_0 + n\mu_s \).

After mixing, the Gibbs free energy of the total system will change:

\[
\Delta G = \Delta E + P\Delta V - T\Delta S,
\]

(20.35)

but if we assume that the solution is ideal, its volume and energy do not depend on the concentration of the solute, so we may assume \( \Delta E \) and \( \Delta V \) are both zero. Therefore,

\[
\Delta G = -T\Delta S.
\]

(20.36)

That is, \( G \) changes only due to the mixing entropy. We model the solution as a lattice gas mixture as illustrated above, so we have

\[
\Delta S = k_B \log \left( \frac{N + n}{n} \right) = -Nk_B \log(1 - x) - nk_B \log x
\]

(20.37)
immediately from Boltzmann’s formula. Therefore, the Gibbs free energy after mixing is,

\[ G = N\mu_0 + n\mu_s + Nk_B T \log(1 - x) + nk_B T \log x. \]  (20.38)

Although we can immediately read off the chemical potentials after mixing, we honestly differentiate \( G \) to get the chemical potentials:

\[ \mu_{\text{solv}} = \left. \frac{\partial G}{\partial N} \right|_{T,P} = \mu_0 + k_B T \log(1 - x). \]  (20.39)

Analogously, we have

\[ \mu_{\text{solute}} = \mu_s + k_B T \log x. \]  (20.40)

### 20.11 Raoult’s law

In the above we interpreted the two components as a solvent and a solute, but we could interpret the mixture as the mixed liquid consisting of two liquids I and II. After mixing, how can we confirm that the chemical potentials of the components can be written as

\[ \begin{align*}
\mu_I &= \mu_{I0} + k_B T \log x_I, \quad (20.41) \\
\mu_{II} &= \mu_{II0} + k_B T \log x_{II}, \quad (20.42)
\end{align*} \]

where \( x_I \) (or \( x_{II} \)) is the mole fraction of component I (II), the temperature is \( T \), and the chemical potential of pure substances (at the pressure we are working) are denoted with suffix 0? We have only to study the partial pressure of the corresponding gas components in the vapor in equilibrium with the mixed liquid. We assume that the vapors are ideal gases and their chemical potentials have the forms of (20.29) or (20.34). Let us write the chemical potentials of individual pure gases at \( T \) and under the atmospheric pressure as \( \mu_{I G}^\ominus, \mu_{II G}^\ominus \). If we write the vapor pressure of pure liquids as \( P_{I0}, P_{II0} \) (in atm), we have\(^{244}\)

\[ \begin{align*}
\mu_{I0} &= \mu_{I G}^\ominus + k_B T \log P_{I0}, \quad (20.43) \\
\mu_{II0} &= \mu_{II G}^\ominus + k_B T \log P_{II0}. \quad (20.44)
\end{align*} \]

\(^{244}\)Here, pressures measured in a particular unit (in atm) appears as it is in the logarithm, so you must respect the unit when you use the formulas.
If the partial pressures in the vapor are $P_I$ and $P_{II}$, then their chemical potentials read $\mu_{IG}$ and $\mu_{IIG}$

$$\begin{align}
\mu_{IG} &= \mu_{IG}^\ominus + k_B T \log P_I, \\
\mu_{IIG} &= \mu_{IIG}^\ominus + k_B T \log P_{II}.
\end{align}$$

(20.45) (20.46)

The coexistence condition of the gas mixture and the liquid mixture is $\mu_I = \mu_{IG}$ and $\mu_{II} = \mu_{IIG}$, so we have

$$\begin{align}
\mu_{IG}^\ominus + k_B T \log P_{I0} + k_B T \log x_I &= \mu_{IG}^\ominus + k_B T \log P_I, \\
\mu_{IIG}^\ominus + k_B T \log P_{II0} + k_B T \log x_{II} &= \mu_{IIG}^\ominus + k_B T \log P_{II}.
\end{align}$$

(20.47) (20.48)

Consequently, we have arrived at Raoult’s law:

$$P_I = x_I P_{I0}, \quad P_{II} = x_{II} P_{II0}.$$  

(20.49)

That is, the partial vapor pressure of a component is identical to its pure vapor pressure \times its mole fraction in the solution.\(^{245}\)

If we combine this and Dalton’s law of partial pressure, we may conclude that the amount of a gas dissolved in a solvent is proportional to the gas pressure (Henry’s law).\(^{246}\)

\section*{20.12 Osmotic pressure}

The osmotic pressure $\pi$ is the required ‘extra’ pressure to prevent flowing in of solvent molecules through the semipermeable membrane that blocks solute molecules (see Fig. 20.5). That is, the pressure of the solution side must be increased as $P \rightarrow P + \pi$. This result is, as we have already seen, an important ingredient of Einstein’s theory of the Brownian motion.\(^{247}\)

\(^{245}\)François-Marie Raoult (1830-1901). He also pointed out the melting point depression for the first time (1878). This was a key to demonstrate that electrolytes indeed dissociate.

\(^{246}\)William Henry (1774-1836). The law was published in 1803. [1803: William Symington demonstrates his Charlotte Dundas, the "first practical steamboat”, in Scotland; Beethoven: Symphony No. 3 "Eroica" (Thielemann, VPO)]

\(^{247}\)There we wished to know the force acting on the solute molecules (suspended particles). We can relate the osmotic pressure and this force as follows. Suppose the pressure is in balance when a mere piston is placed between the solution and the pure solvent. Now, we replace the piston with a rigid membrane that blocks only solute molecules. This means the force (pressure) due to the solute particles is sustained by the membrane. Thus, the pressure contributed by the solvent to the outside of the solution is reduced by this pressure which is equal to $\pi$, allowing the solvent molecules to flow in from the pure solvent side. That is, the force corresponding to $\pi$ is acting on the solute molecules.
Since both sides of the membrane must have the same chemical potentials for the solvent molecules, because they can go through the membrane (see (20.39)):

$$\mu_{\text{solv}}(P + \pi, T) = \mu_0(P + \pi, T) + k_B T \log(1 - x) = \mu_0(P, T), \quad (20.50)$$

so

$$-\mu_0(P + \pi, T) + \mu_0(P, T) = k_B T \log(1 - x) \simeq -k_B T x. \quad (20.51)$$

As we have already seen in the above $d\mu = (V/N)dP$ under constant temperature, so the above equation reads (Taylor expansion!)

$$-(V/N)\pi = -k_B T (n/N) \Rightarrow \pi V = n k_B T. \quad (20.52)$$

This is called van’t Hoff’s law, which Einstein used in his Brownian motion theory (actually, he derived this equation by himself).

### 20.13 Colligative properties

A very similar question is the melting-point depression or the boiling point elevation due to solute (see the problems at the end of his lecture).

Raoult’s law, van’t Hoff’s law, Henry’s law, boiling-point elevation, melting-point depression, etc. are all independent of the chemical nature of the solute and are due to the $\log(1 - x)$ or $\log x$ term in the chemical potential (i.e., only due to the particle number ratio), so we may understand them in a unified fashion. Therefore, these phenomena are traditionally said to exhibit the colligative properties.

### 20.14 Chemical reactions

Here, an elementary exposition of equilibrium chemical reactions is given. Without (irreversible) chemical reactions no atomism was possible. Furthermore, the idea of

\[248\text{ligated together}\]
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.

Since the general formulas may be cumbersome, in this lecture, we use the following reaction to illustrate the general formulas:

\[ \text{N}_2 + 3\text{H}_2 \leftrightarrow 2\text{NH}_3. \]  

This formula implies that one molecule (or one mole\(^{249}\)) 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 an appropriate set of elementary reactions.\(^{250}\)

The left hand side of (20.53) 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

\[ -\text{N}_2 - 3\text{H}_2 + 2\text{NH}_3 = 0. \]

Thus, generally any reaction may be written as

\[ \sum \nu_i C_i = 0, \]

where \(\nu_i\) are signed stoichiometric coefficients for chemical \(C_i\); \(\nu_i > 0\) (resp., \(\nu_i < 0\)) implies \(i\) is a product (resp., a reactant).

### 20.15 Activity of Chemical

For a gas mixture with the partial pressure \(P_i\) of chemical \(i\), we may write its chemical potential per mole\(^{251}\) as

\[ \mu_i = \mu_i^\ominus + RT \log P_i. \]

\(^{249}\)The quantity of one mole (mole number) is, as already noted, defined by the same number (Avogadro’s number) of molecules as contained in 12 g of \(^{12}\text{C}\).

\(^{250}\)As the actual elementary reactions in the gas phase, such a reaction as (20.53) is very unusual, because elementary reactions are unimolecular decay or binary collision type reactions.

\(^{251}\)Up to this point we studied everything per molecule, so \(k_B\) appeared in the expression of chemical potentials, but in the chemical reaction part of this lecture, the chemical potentials per mole will be used, so \(k_B\) is everywhere replaced with \(R\).
Here, $\mu_i^\oplus$ is the chemical potential for, e.g., $P_i = 1$ (in, say, MPa, atm, etc.\textsuperscript{252}). In solutions the chemical potential of a solute $i$ in a solution is written as

$$\mu_i = \mu_i^\oplus(T, P) + RT \log a_i, \quad (20.57)$$

where $a_i$ is called the activity of chemical $i$, which is close to the mole fraction $x_i$ when the solution is dilute.

### 20.16 Equilibrium condition for reactions: the law of mass action

The equilibrium condition for the reaction (20.55) reads

$$0 = \sum_i \nu_i \mu_i = \sum_i \nu_i [\mu_i^\oplus(T, P) + RT \log a_i]. \quad (20.58)$$

Or, (assuming the constant $TP$ condition)

$$-\Delta G^\ominus \equiv -\sum_i \nu_i \mu_i^\oplus(T, P) = RT \log \left( \prod_i a_i^{\nu_i} \right). \quad (20.59)$$

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

$$K(T, P) = e^{-\Delta G^\ominus/RT} = \frac{\cdots a_p^{\nu_p} \cdots}{\cdots a_r^{-\nu_r} \cdots}, \quad (20.60)$$

where the numerator have all the products, and the denominator all the reactants. (20.60) 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). The equilibrium constant for the reaction (20.53) is given by

$$K(T, P) = \frac{[\text{NH}_3]^2}{[\text{N}_2][\text{H}_2]^3}. \quad (20.61)$$

Here, $[X]$ generally describes the partial pressure (fugacity) of chemical $X$ in the gas phase reaction or the molarity or mole fraction (or activity) in the solution.

In principle, the chemical equilibrium constant may be statistical-mechanically computed. However, except for ideal gasses, it is prohibitively hard to compute the needed chemical potentials, so, for almost all interesting examples of chemical reactions, theoretical calculations are useless.

---

\textsuperscript{252}Since a dimensional quantity appears in the logarithm, when you use such formulas, you must stick to the unit being used.
20.17 Shift of chemical equilibrium

If we differentiate the equilibrium constant with respect to $T$, we can obtain the heat of reaction, that is, $\Delta H$ (enthalpy change) due to reaction. The Gibbs-Helmholtz relation (or its analog for the Gibbs free energy) tells us

$$\frac{\partial \log K}{\partial T} \bigg|_P = \frac{\Delta H^{\circ}}{RT^2}, \quad (20.62)$$

where $\Delta H^{\circ}$ is the enthalpy change for the ‘standard state.’ This is called van’t Hoff’s equation. Similarly,

$$\frac{\partial \log K}{\partial P} \bigg|_T = -\frac{\Delta V^{\circ}}{RT}, \quad (20.63)$$

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).

(20.62) 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 shift back from the product system to the original reactant system). This is an example of Le Chatelier’s principle (Lecture 17) asserting that “the response to a perturbation is in the direction to reduce its effect.” (20.63) is also its example. Needless to say, these are manifestations of the stability of our world.
Q20-1 [Gibbs relation for densities]
Show that the Gibbs relation for densities (extensive quantities per volume): \( e = E/V \) (internal energy density), \( s = S/V \) (entropy density), \( n = N/V \) (number density), \( \bar{x} = X/V \), etc. reads

\[
de = Tds + \mu dn + xd\bar{x}. \tag{20.64}
\]

**Soln.**

From (20.9), we get

\[
e = Ts - P + \mu n + x\bar{x}. \tag{20.65}
\]

On the hand, dividing the Gibbs-Duhem relation (20.11) with \( V \), we get

\[
svd\bar{T} - dP + \eta dn + \bar{x}dx = 0 \tag{20.66}
\]

Therefore, differentiating (20.65) and using (20.66), we get

\[
de = Tds + \mu dn + xd\bar{x}. \tag{20.67}
\]

How about the Gibbs relation per unit mass or per mole?

Q20-2 Let us continue the lattice gas example to understand the effect of impurities on the phase transition temperatures.

![Lattice gas model](image)

We know the chemical potential of the solvent molecules (majority) reads

\[
\mu_L = \mu_L^\oplus + k_B T \log(1 - x). \tag{20.68}
\]

Here, \( x \) is the mole fraction of the solute molecules (blue particles in Fig. 20.6) the solution. \( \mu_L^\oplus \) is the chemical potential of the pure solvent liquid. Let us write the chemical potential of the pure solid phase of the solvent as \( \mu_{S}^\oplus \).

(1) If we cool the solution, a solid phase of the solvent molecules emerges. When solidification occurs, impurity molecules (i.e., solute molecules) are largely excluded.
from the emerging solid. Let us idealize (not a bad approximation) the solid phase to be pure. Let $T_m$ be the melting point of the pure substance. This implies

$$\mu_L(T_m, P) = \mu_S(T_m, P). \quad (20.69)$$

What is the equilibrium coexistence temperature of the pure solid and the solution with a mole fraction $x$ of impurity molecules? You have only to compute the melting temperature shift $\Delta T$ to order $x$. Assume that the latent heat of melting is $L$ (per molecule).

(2) We may assume that the vapor of the solvent at $T$ may be approximated as an ideal gas. At $T$ its pressure is $P$. Now, we add the solute molecules as impurity to this solvent. You may assume that the solute molecules cannot escape the liquid phase, so the vapor phase still consists of pure solvent molecules. What is the pressure change $\Delta P$ due to the addition of the mole fraction $x$ of the impurity molecules?

(3) You must have obtained $\Delta P/P = -x$ from (2), assuming the gas volume is much larger than the liquid volume so the latter may be ignored. $|\Delta P|$ is 2850 Pa, when 23.3 g of a substance is solved in 100 g of water at 100 °C. What is the molecular weight of this substance?

**Soln.**

(1) Let us assume that at temperature $T_m + \Delta T$ the equilibrium between the solvent crystal and the solution holds:

$$\mu_S(T_m + \Delta T, P) = \mu_L(T_m + \Delta T, P) + k_B(T_m = \Delta T) \log(1 - x). \quad (20.70)$$

That is, to order $x$, we may Taylor-expand the above condition as

$$\frac{\partial [\mu_S(T_m, P) - \mu_L(T_m, P)]}{\partial T_m} \Delta T = -k_B T_m x \quad (20.71)$$

Denoting extensive quantities per molecule with lower case letters corresponding to the standard notation, Gibbs’ relation gives $d\mu = -s dT + v dP$, so the above equation reads (the ‘pure sign’ $\ominus$ is omitted)

$$[s_L(T_m, P) - s_S(T_m, P)] \Delta T = -k_B T_m x \quad (20.72)$$

From the latent heat $L$

$$s_L(T_m, P) - s_S(T_m, P) = L/T_m \quad (20.73)$$

Thus, we have reached

$$\Delta T = -\frac{k_B T_m^2}{L} x < 0 \quad (20.74)$$
That is, the melting point is lowered by the amount proportional to the impurity concentration. This is called the *melting-point depression*.

(2) Let us denote the ideal gas chemical potential as \( \mu_G(T, P) \). From the general form, we may write

\[
\mu_G(T, P) = \mu_G^\circ(T) + k_B T \log P. \tag{20.75}
\]

Therefore, the equilibrium between the pure solvent and its vapor requires at pressure \( P \)

\[
\mu_L^\circ(T, P) = \mu_G^\circ(T) + k_B T \log P, \tag{20.76}
\]

and the equilibrium between the solution and its vapor requires at pressure \( P + \Delta P \)

\[
\mu_L^\circ(T, P + \Delta P) + k_B T \log(1 - x) = \mu_G^\circ(T) + k_B T \log(1 + \Delta P/P). \tag{20.77}
\]

Subtracting (20.76) from (20.77), we obtain

\[
\mu_L^\circ(T, P + \Delta P) - \mu_L^\circ(T, P) + k_B T \log(1 - x) = k_B T \log(1 + \Delta P/P). \tag{20.78}
\]

Taylor-expanding this to order \( x \), we get

\[
\left. \frac{\partial \mu_L}{\partial P} \right|_T \Delta P - k_B T x P = \frac{k_B T}{P} \Delta P. \tag{20.79}
\]

The partial derivative here gives the volume (per molecule) of the solvent liquid, which may be neglected relative to the gas volume \( k_B T/P \). After this approximation we get the famous equation

\[
\Delta P = -x P. \tag{20.80}
\]

(3) Let \( M \) be the molecular weight of the solute. Then, since the ambient pressure is 1 atm (as seen from the boiling point of the water)

\[
x = \frac{23.3/M}{100/18 + 23.3/M} = \frac{|\Delta P|}{|\Delta P|/P} \approx \frac{100}{18} \frac{|\Delta P|}{P} = \frac{100}{18} \frac{2850}{1.013 \times 10^5}. \tag{20.81}
\]

or

\[
\frac{23.3}{M} \left( \frac{100/18}{|\Delta P|/P} \right) \simeq \frac{100}{18} \frac{|\Delta P|}{P} = \frac{100}{18} \frac{2850}{1.013 \times 10^5}. \tag{20.82}
\]

Hence, \( M \approx 149 \).

**Q20-3.** Let us consider the reaction to synthesize ammonia:

\[
\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_3. \tag{20.83}
\]
Its equilibrium constant may be written in terms of partial pressures as

$$K = \frac{P^2_{\text{NH}_3}}{P_{\text{N}_2}P^3_{\text{H}_2}} = 1.5 \times 10^{-5}, \quad (20.84)$$

if the (partial) pressures are in atm (at 500 °C).

(1) If you wish to synthesize ammonia, is it more or less advantageous to increase the total pressure of the reaction vessel? You must justify your answer (perhaps quoting the relevant equation(s)).

(2) Suppose the atomic ratio of $N$ and $H$ is 1:3 (i.e., the stoichiometric ratio). If 90% of reactants are converted into ammonia in equilibrium, what is the total pressure $P$ of the mixture? You may treat the gases as ideal gases.

**Soln.**

(1) This is le Chatelier. Since the volume decreases if the reaction proceeds, increasing pressure should shift the reaction to the ammonia side. Or more precisely, we use

$$\frac{\partial \log K}{\partial P} \bigg|_T = -\frac{\Delta V}{RT}, \quad (20.85)$$

where $\Delta V$ is the volume change due to the reaction. In this case it is negative, so increasing $P$ increases $K$, that is, the reaction shifts two the right.

(2) We need the final mole fractions of the chemicals. If 100$x\%$ of $\text{N}_2$ has been converted into ammonia and there was 1 mole of nitrogen gas, the total number of molecules in moles is

$$1 - x + 3(1 - x) + 2x = 4 - 2x. \quad (20.86)$$

Therefore, the partial pressures after equilibration read

$$P_{\text{N}_2} = P \frac{1 - x}{4 - 2x}, P^3_{\text{H}_2} = P \frac{3 - 3x}{4 - 2x}, P^2_{\text{NH}_3} = P \frac{2x}{4 - 2x}. \quad (20.87)$$

Hence,

$$K = \frac{[2x/(4 - 2x)]^2}{P^2[(1 - x)/(4 - 2x)][(3 - 3x)/(4 - 2x)]^3} = \frac{16x^2(2 - x)^2}{27P^2(1 - x)^4} \quad (20.88)$$

For $x = 0.9$

$$P^2 = 5808/K = 3.87 \times 10^8, \quad (20.89)$$

$$P = 1.97 \times 10^4 \text{ atm.}$$
21 Grand canonical ensemble and ideal quantum systems

Summary
* Grand canonical partition function/ensemble is introduced.
* Noninteracting quantum systems are discussed with the aid of grand partition function.
* Pressures of fermions and bosons are compared.
* $PV = 2E/3$ for ‘any’ ideal gas.

Key words
grand canonical partition function, grand canonical ensemble, fermion, boson, Bose-Einstein distribution, Fermi-Dirac distribution

What you should be able to do
* Be able to explain why $PV$ is directly obtained from the grand canonical partition function.
* Intuitively understand low temperature noninteracting fermion and boson systems.
* In particular, compare the pressures of boson and fermion systems intuitively.
* Be able to derive the one particle density of states $D_t(\varepsilon)$. Note $D_t \propto \sqrt{\varepsilon}$.

21.1 Grand canonical ensemble/partition function: motivation
We now introduce another ensemble. We know microcanonical, canonical, and generalized canonical ensembles (e.g., pressure ensembles). The fourth law 9.4 tells us

$$A = E - TS = -PV + \mu N. \quad (21.1)$$

Can you directly derive this, following the logic used in Lecture 20 to demonstrate the Gibbs-Duhem relation (i.e., mimic the fine letters in 20.5)? Therefore, (21.1) implies

$$PV = -A + \mu N = ST - E + \mu N, \quad (21.2)$$

That is,

$$\frac{PV}{T} = S - \frac{E}{T} + \frac{\mu}{T}N. \quad (21.3)$$

This is a Legendre transformation of entropy, so there must be an ensemble that directly gives $PV/T$ or $PV/k_B T$. 299
Compare the following formulas:

\[ S = k_B \log w(E, V, X), \quad (21.4) \]
\[ -\frac{A}{T} = S - \frac{E}{T} = k_B \log Z(T, V, X). \quad (21.5) \]

We know with the aid of Boltzmann’s principle

\[ Z(T, V, X) = \int dE w(E, V, X) e^{-E/k_BT} = \int dE e^{[S(E) - E/T]/k_B}. \quad (21.6) \]

Thus, we can easily mimic this to get

\[ \frac{PV}{T} = S - \frac{E - \mu N}{T} = k_B \log \Xi(T, V, \mu) \quad (21.7) \]

with

\[ \Xi(T, V, \mu) = \int dE \sum_N w(E, V, N) e^{-(E-\mu N)/k_BT} = \sum_{N=0}^{\infty} Z(T, V, N)e^{\beta\mu N}. \quad (21.8) \]

\( \Xi \) is called the grand (canonical) partition function, which describes the system thermostatted and chemostatted with a reservoir at temperature \( T \) and chemical potential \( \mu \). Recall that \( \mu \) is the needed work to push one molecule into the system, so by adjusting \( \mu \) you can regulate the average number of particles in the system. If a system is macroscopic, the fluctuation of the total number of particles is irrelevant, so you may use \( \mu \) to control \( N \).

### 21.2 Grand canonical ensemble/partition function: summary

Let us summarize:

\[ \frac{PV}{T} = -\frac{A}{T} + \frac{\mu N}{T} = k_B \log \Xi(T, V, \mu) \quad (21.9) \]

with

\[ \Xi(T, V, \mu) = \sum_{N=0}^{\infty} Z(T, V, N)e^{\beta\mu N} = \sum_{\text{microstates}} e^{\beta(H-\mu N)}, \quad (21.10) \]

where the summation over ‘microstates’ means all the possible microstates allowed to the system irrespective of the total number of particles in the system.

The ensemble equivalence holds here as well: If \( N \gg \log N \), then you can use any ensemble you wish. For example, if you have about a few thousand particles confined
in a trap, you may use the grand canonical formalism above to describe the system. The Gibbs relation reads
\[ d \left( \frac{PV}{T} \right) = -E d\frac{1}{T} + P dV + N d\mu. \quad (21.11) \]
This has a (statistical-mechanically) more convenient form:
\[ d \log \Xi = -E d\beta + \beta P dV + N d(\beta \mu). \quad (21.12) \]

### 21.3 Example: adsorption

Let us solve an example problem (which is made deliberately slightly complicated) to become familiar with the use of chemical potentials and the grand canonical formalism. Suppose there is a gas mixture consisting of two distinct molecular species $A$ and $B$. The mixture is an ideal gas and the partial pressure of $X$ is $P_X$ ($X = A$ or $B$). The gas is in equilibrium with an adsorbing metal surface on which there are $N$ adsorption sites. Molecule $X$ adsorbed at a site is with energy $\varepsilon_X$ (which is often negative) relative to the one in the gas phase, where $X = A$ or $B$. Each surface site can accommodate at most one $A$ molecule, and at most two $B$ molecules. One adsorbed $A$ atom has 2 different (internal) states (with the same energy), and one adsorbed $B$ molecule has 1 state, but if two $B$ molecules are adsorbed to the same site, then they can together have 5 states with the same energy (the caption of Fig. 21.1 summarizes the system). We wish to know the surface concentration of the atoms when the surface is in equilibrium with the gas mixture. You may assume the gas phase is huge, so you need not worry about its composition change due to adsorption. That is, the gas phase is a chemical reservoir.

We wish to know the average number of $A$ and $B$ atoms on the metal surface. Since we do not know how many particles are on the surface, it must be convenient to use the grand canonical ensemble. Assuming the chemical potentials of $A$ and $B$ as $\mu_A$ and $\mu_B$, respectively, write down the partition function of the metal surface:
\[ \Xi = [1 + 2e^{-\beta(\varepsilon_A - \mu_A)} + e^{-\beta(\varepsilon_B - \mu_B)} + 5e^{-2\beta(\varepsilon_B - \mu_B)}]^N. \quad (21.13) \]
The needed chemical potentials can be computed with the aid of the ideal gas statistical mechanics as we did in the preceding lecture. We have done that calculation, so let us copy the needed results:
\[ \mu_A = k_B T \log(\beta P_A/n_{QA}), \quad \mu_B = k_B T \log(\beta P_B/n_{QB}). \quad (21.14) \]
Figure 21.1: Adsorption of gas particles on a metal surface: $A$: green (2 internal states when adsorbed), $B$: red (if singly adsorbed, with single internal state; if doubly adsorbed, with 5 internal states).

Here, $n_{QX}$ is the ‘quantum density’ depending on $T$ and the mass (see above (20.32)), and $P_{QX} = n_{QX}k_BT$ may be called the ‘quantum pressure.’ We know (we ignore the volume (or the area) change)

$$d\log \Xi = -Ed\beta + N_Ad(\beta\mu_A) + N_Bd(\beta\mu_B),$$

so we obtain

$$N_A = N\frac{2e^{-\beta(\epsilon_A-\mu_A)}}{1 + 2e^{-\beta(\epsilon_A-\mu_A)} + e^{-\beta(\epsilon_B-\mu_B)} + 5e^{-2\beta(\epsilon_B-\mu_B)}}$$

and

$$N_B = N\frac{e^{-\beta(\epsilon_B-\mu_B)} + 10e^{-2\beta(\epsilon_B-\mu_B)}}{1 + 2e^{-\beta(\epsilon_A-\mu_A)} + e^{-\beta(\epsilon_B-\mu_B)} + 5e^{-2\beta(\epsilon_B-\mu_B)}}.$$  

21.4 Microstates for non-interacting indistinguishable particle systems

Let us consider a system consisting of non-interacting particles. Suppose the states of a single particle are numbered as $i = 1, 2, \cdots$. If we assume that all the particles are indistinguishable, then to specify a microstate of a system consisting of such particles, we have only to count the number $n_i$ of particles in the $i$-th one particle state ($n_i$ is also called the occupation number of the $i$th one particle state; do NOT confuse the microstates of the whole system and the one particle states). Or, we have only to make a table of the occupation numbers $\{n_1, n_2, \cdots\}$; we may identify this table and the microstate.

To study the thermodynamics of such a system, we should use the grand canonical ensemble, because we have not specified the total number of particles.
21.5 Grand partition function of indistinguishable particle system

Let \( \varepsilon_i \) be the energy of the \( i \)-th one particle state. The total energy \( \mathcal{E} \) and the total number of particles \( N \) of the microstate \( \{n_1, n_2, \cdots \} \) can be written as

\[
\mathcal{E} = \sum_{i=1} \varepsilon_i n_i, \quad (21.18)
\]

and

\[
N = \sum_{i=1} n_i. \quad (21.19)
\]

Then, the grand canonical partition function must be

\[
\Xi(\beta, \mu) = \sum_{n_1, n_2, \cdots} e^{-\beta \mathcal{E} + \beta \mu N}. \quad (21.20)
\]

Using the microscopic descriptions of \( \mathcal{E} \) and \( N \) ((21.18) and (21.19)), we can rearrange the summation as

\[
\Xi = \prod_i \Xi_i, \quad (21.21)
\]

where

\[
\Xi_i \equiv \sum_{n_i} \exp[-\beta(\varepsilon - \mu)n_i]. \quad (21.22)
\]

This quantity may be called the grand canonical partition function for the \( i \)th one particle state.

21.6 Boson and fermions

In the world it seems that there are only two kinds of particles:

- **bosons**: there is no upper bound for the occupation number;
- **fermions**: the occupation number can be at most 1 (the Pauli exclusion principle).

This is an empirical fact. Electrons, protons, \(^3\text{He}\), etc., are fermions. Mesons, \(^4\text{He}\), \(D\), etc., are bosons.

There is the so-called spin-statistics relation that the particles with half odd integer spins are fermions, and those with integer spins are bosons. The rule applies also to compound particles such as hydrogen atoms. Thus, \(H\) and \(T\) are bosons, but their nuclei are fermions. \(D\) and \(^3\text{He}\) are fermions. \(^4\text{He}\) is a boson, and so is its nucleus.

For a neutral system consisting of \(+\) and \(−\) charged particles (e.g., the usual
electron-nucleus system) it is proved that at least $+$ or $-$ species must be all fermions for the system to be stable. Here, ‘stable’ means that there is a positive number $B$ such that the system energy $E$ satisfies $E > -NB$, where $N$ is the number of particles in the system. That is, for the world to be stable, we desperately need fermions.

### 21.7 Ideal boson systems

For bosons, any number of particles can occupy the same one particle state, so the occupation number of a particular one particle state can be any of $0, 1, 2, \cdots$. Therefore,

$$\Xi_i = \sum_{n=0}^{\infty} e^{-\beta (\varepsilon_i - \mu) n} = \left( 1 - e^{-\beta (\varepsilon_i - \mu)} \right)^{-1}. \quad \text{(21.23)}$$

The mean occupation number of the $i$-th state is given by

$$\langle n_i \rangle = \sum_{n=0}^{\infty} n_i e^{-\beta (\varepsilon_i - \mu) n} / \Xi_i, \quad \text{(21.24)}$$

so we conclude

$$\langle n_i \rangle = \left. \frac{\partial \log \Xi_i}{\partial \beta \mu} \right|_{\beta} = k_B T \left. \frac{\partial \log \Xi_i}{\partial \mu} \right|_{T} = \frac{1}{e^{\beta (\varepsilon_i - \mu)} - 1}. \quad \text{(21.25)}$$

This distribution is called the **Bose-Einstein distribution**.

If the (one-particle) ground-state energy is zero, then the ground state occupancy is

$$\langle n_{\text{ground}} \rangle = \frac{1}{e^{-\beta \mu} - 1}, \quad \text{(21.26)}$$

but this should not be negative, so $\mu \leq 0$. That is, notice that the chemical potential must be smaller than the (one-particle) ground state energy to maintain the positivity of the average occupation number.

### 21.8 Ideal fermion systems

For fermions, at most one particle can occupy the same one particle state, the occupation number of a particular one particle state is 0 or 1. Therefore,

$$\Xi_i = \sum_{n=0}^{1} e^{-\beta (\varepsilon_i - \mu) n} = 1 + e^{-\beta (\varepsilon_i - \mu)}. \quad \text{(21.27)}$$
The mean occupation number of the $i$-th state is given by

$$
\langle n_i \rangle = \sum_{n=0}^{1} n_i e^{-\beta(\epsilon_i - \mu)n}/\Xi_i,
$$

so we conclude

$$
\langle n_i \rangle = \frac{\partial \log \Xi_i}{\partial \beta \mu} \bigg|_{\beta} = k_B T \left. \frac{\partial \log \Xi_i}{\partial \mu} \right|_T = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1}.
$$

This distribution is called the Fermi-Dirac distribution.

Notice that usually the ground state energy is chosen as the origin of energy, so

$$
\langle n_{\text{ground}} \rangle = \frac{1}{e^{-\beta \mu} + 1}.
$$

It is important to recognize the qualitative features of this Fermi-Dirac distribution function (see Fig. 21.2). The distribution has a cliff of width of order $k_B T$. In the $T \to 0$ limit, it has a vertical cliff at $\epsilon = \mu$, which is called the Fermi level. Notice that $\mu > 0$ is required, if the temperature is low enough.

Figure 21.2: The cliff has a width of order $k_B T$. $\mu$ is called the Fermi level. The symmetry noted in the figure is the so-called particle-hole symmetry.

### 21.9 Classical limit

The distribution functions of the occupation numbers are quite different from the classical distribution function obtained by Maxwell and Boltzmann. The difference should be due to the quantum interference among particles (or particle wave functions) when the number density is not low. Therefore, in order to obtain the classical

---

253Do not forget that $\mu$ for a fixed $N$ is temperature dependent. The energy of the highest occupied state at $T = 0$ is called the Fermi energy.
limit, we must take the occupation number 0 limit to avoid quantum interference among particles. The chemical potential $\mu$ is a measure of the “strength” of the chemostat to push particles into the system. Thus, we must make the chemical potential extremely small: $\mu \rightarrow -\infty$.

In this limit both Bose-Einstein (21.26) and Fermi-Dirac distributions (21.30) reduce to the Maxwell-Boltzmann distribution as expected:

$$\langle n_i \rangle \rightarrow N e^{-\beta \varepsilon_i},$$  \hspace{1cm} (21.31)

where $N = e^{\beta \mu}$ is the normalization constant determined by the total number of particles in the system.

Notice that $\mu \rightarrow -\infty$ is far away from the situations where quantum effects are important; $\mu \approx 0$ for bosons and $\mu > 0$ for fermions.

### 21.10 Intuitive pictures

Before going to the equations of state, let us try to build our intuition. Suppose there are only three one-particle states with energies 0, $\varepsilon$ and $3\varepsilon$, and there are three particles. Make a table of all the microstates of the three-particle system for bosons and for fermions.

<table>
<thead>
<tr>
<th>total energy</th>
<th>fermion case</th>
<th>boson case</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3\varepsilon$</td>
<td>4$\varepsilon$ 0 $3\varepsilon$ 2$\varepsilon$ 4$\varepsilon$ 6$\varepsilon$ 3$\varepsilon$ 5$\varepsilon$ 7$\varepsilon$ 9$\varepsilon$</td>
<td>$$\bullet$$ $$\bullet$$ $$\bullet$$ $$\bullet$$ $$\bullet$$ $$\bullet$$ $$\bullet$$ $$\bullet$$ $$\bullet$$ $\bullet$</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>$$\bullet$$</td>
<td>$$\bullet$$ $\bullet$ $\bullet$ $\bullet$ $\bullet$ $\bullet$ $\bullet$</td>
</tr>
<tr>
<td>0</td>
<td>$$\bullet$$</td>
<td>$$\bullet$$ $\bullet$ $\bullet$ $\bullet$ $\bullet$ $\bullet$</td>
</tr>
</tbody>
</table>

Figure 21.3: identical particles in three one particle states; fermion (leftmost) and boson cases.

Fermions first:

<table>
<thead>
<tr>
<th>microstate</th>
<th>0</th>
<th>$\varepsilon$</th>
<th>3$\varepsilon$</th>
<th>total energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4$\varepsilon$</td>
</tr>
</tbody>
</table>
Bosons:

<table>
<thead>
<tr>
<th>microstate</th>
<th>0</th>
<th>ε</th>
<th>3ε</th>
<th>total energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>ε</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3ε</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2ε</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4ε</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>6ε</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3ε</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>5ε</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7ε</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>9ε</td>
</tr>
</tbody>
</table>

Suppose there are 100 identical spinless\textsuperscript{254} bosons or fermions whose \( s \)-th one-particle state has an energy \( \varepsilon_s = s \varepsilon (s \in \mathbb{N}) \). These particles do not interact. For the boson case, at \( T = 0 \) all the particles are in the lowest energy one-particle state (see Fig. 21.4). For fermions, all the low-lying one particle states are completely filled up to some energy level that corresponds to \( \mu \). Notice that the ground state energy of the fermion and boson systems are quite different.

The low-lying excited microstates are also in Fig. 21.4 (right). For the boson case all the particles have equal chance to be excited, but in the case of fermions, only the particles near the Fermi level can be excited (and excited particles leave holes). This should tell you something about the specific heat of these systems (later).

\textsuperscript{254}Spinless’ implies that these particles do not have any internal degrees of freedom.
21.11 Pressure of ideal systems
The distinction between fermions and bosons show up clearly in pressure:

\[
\frac{PV}{k_B T} = \log \Xi = \mp \sum_i \log \left(1 \mp e^{-\beta (\varepsilon_i - \mu)} \right).
\] (21.32)

If \(T, V, \mu\) are the same, then the pressure of the system consisting of the particles with the same single-particle energy states (i.e., the the same density of states) shows the following ordering (BE = Bose-Einstein, MB = Maxwell-Boltzmann, FD = Fermi-Dirac):

\[P_{BE} > P_{MB} > P_{FD}.\] (21.33)

To see \(P_{BE} > P_{FD}\) we use (21.32). Let \(x = e^{-\beta (\varepsilon_i - \mu)}\) and we compare \(-\log(1 - x)\) and \(\log(1 + x)\) (see Fig. 21.5). This figure must be enough. This is also easily seen from \(1/(1 - x) = 1 + x + x^2 + \cdots > 1 + x\) for \(x \in (0, 1)\).

Figure 21.5: \(\log(1 + x)\) is always below \(x\). Now, \(-\log(1 - x)\) is obtained from \(\log(1 + x)\) by making mirror images \(x \rightarrow -x\) and then \(y \rightarrow -y\). Obviously, \(x < -\log(1 - x)\).

21.12 Pressure under constant \(N\) (the usual case)
In contrast, if \(T, V, N\) are the same (the usually more interesting case than the above case), then (usually \(P_{FD} \gg P_{MB}\))

\[P_{FD} > P_{MB} > P_{BE}.\] (21.34)

To show this requires some trick, so a formal demonstration is below with fine letters, but intuitively this can be understood by the extent of effective particle-particle attraction as is illustrated in Fig. 21.6. The figure not only suggests the pressures, but also suggests the extent of particle density fluctuations. The particle density fluctuations in a boson system are larger than those in a fermion system.
2/3
2/4=1/2
0

Figure 21.6: Two-particle two box illustration of statistics. The fractions in the right denote the relative weights of the states for which effective attraction can be seen. (BE = Bose-Einstein, MB = Maxwell-Boltzmann, FD = Fermi-Dirac)

As seen from the ‘effective attraction weights’ in the above figure, the fermion system exhibits the largest pressure; fermions avoid each other (Pauli’s exclusion principle), so they hit the wall more often

$\log \Xi_{FD} > \langle N \rangle > \log \Xi_{BE}$.  

Let us see the first inequality:\textsuperscript{255} Writing $x_j = e^{-\beta (\varepsilon_j - \mu)}$, we have

$$\log \Xi_{FD} - \langle N \rangle = \sum_j \left[ \log(1 + x_j) - \frac{x_j}{1 + x_j} \right].$$

We are done, because for $x > 0$\textsuperscript{256}

$$\log(1 + x) - \frac{x}{1 + x} > 0.$$

Similarly, we can prove the second inequality in (21.35).

\textbf{21.13 Universal $P$-$E$ relation: introduction}

Let $D(\varepsilon)d\varepsilon$ denote the number of single particle states whose energy is between $\varepsilon$ and $\varepsilon + d\varepsilon$. $D(\varepsilon)$ is called the one-particle state density (or density of states of the one-particle system). If we know this, the pressure (21.32) can be rewritten as

$$PV = \mp k_B T \int d\varepsilon \ D(\varepsilon) \log \left( 1 \mp e^{-\beta (\varepsilon - \mu)} \right).$$

\textsuperscript{255}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.

\textsuperscript{256}Consider the derivatives.
In the present case $D(\varepsilon)$ is the density of states for a particle confined in a 3D box of volume $V$, which we will denote by $D_t(\varepsilon)$, where the suffix $t$ implies the translational degrees of freedom.

We know for a classical ideal gas

$$PV = \frac{2}{3}E,$$  

where $E$ is the internal energy. ‘Miraculously,’ this is true for non-interacting fermions and bosons (and their mixtures as well). The formulas for $PV$ and $E$ are quite different from the classical case. Here,

$$E = \int d\varepsilon D_t(\varepsilon)\varepsilon \langle n(\varepsilon) \rangle = \int d\varepsilon D_t(\varepsilon) \frac{\varepsilon}{e^{\beta(\varepsilon - \mu)} + 1}. \quad (21.40)$$

### 21.14 Density of one-particle states

To demonstrate this, we need $D_t(\varepsilon)$. First, a quick way is explained to derive the formula appropriate for a statistical mechanics course. We count the number of microscopic states for a single particle up to some energy $\varepsilon$. To do this we use the classical-quantum mechanics correspondence: the number of quantum states in the phase volume element $dpdq$ is given by $dpdq/h^3$. Then, the total number of quantum states for a single particle whose energy is less than or equal to $\varepsilon$ must be given by

$$\frac{1}{h^3} \int_{q \in V} dq \int_{|p| < \sqrt{2m\varepsilon}} dp = \int_0^\varepsilon d\varepsilon D_t(\varepsilon), \quad (21.41)$$

that is,

$$\int_0^\varepsilon d\varepsilon D_t(\varepsilon) = \frac{4\pi}{h^3} V \int_0^{\sqrt{2m\varepsilon}} p^2 dp. \quad (21.42)$$

Differentiating this with $\varepsilon$, we get

$$D_t(\varepsilon) = \frac{4\pi}{h^3} V \sqrt{2m} (\sqrt{2m\varepsilon})^2 = 2\pi V \left(\frac{2m}{h^2}\right)^{3/2} \varepsilon^{1/2}. \quad (21.43)$$

You can easily extend this approach to higher or lower dimensional spaces, and to the cases with other $p-\varepsilon$ relations (dispersion relations). $D_t(\varepsilon) \propto \varepsilon^{1/2}$ is important (worth memorizing) for the case with $\varepsilon \propto p^2$ in 3-space.

That is, $D_t(\varepsilon) = \gamma V \varepsilon^{1/2}$, where $\gamma = 2\pi(2m/h^2)^{3/2}$ is a constant. This can be
obtained by a dimensional analytical idea as well. $D_t(\varepsilon)h^3d\varepsilon$ must have the dimension of the phase volume whose dimension is $L^3$ times [momentum]$^3$, so $D_t(\varepsilon)h^3 \propto V\sqrt{\varepsilon^3/\varepsilon} = V\varepsilon^{1/2}$.

Note the following relation that can easily be seen from $\int \varepsilon^{1/2}d\varepsilon = (2/3)\varepsilon^{3/2} = (2\varepsilon/3)\varepsilon^{1/2}$

$$\int_0^\varepsilon d\varepsilon D_t(\varepsilon) = \frac{2}{3}\varepsilon D_t(\varepsilon). \quad (21.44)$$

### 21.15 Universal $P$-$E$ relation: demonstration

Let us return to our problem. The pressure can be rewritten as (the fundamental theorem of calculus)

$$PV = \pm k_B T \int d\varepsilon \left[\int_0^\varepsilon d\varepsilon' D_t(\varepsilon')\right]' \log \left(1 \mp e^{-\beta(\varepsilon-\mu)}\right). \quad (21.45)$$

Performing an integration by parts, we get

$$PV = \pm k_B T \left[\int_0^\varepsilon d\varepsilon' D_t(\varepsilon') \log \left(1 \mp e^{-\beta(\varepsilon-\mu)}\right)\right]_0^\infty \pm k_B T \int d\varepsilon \left[\int_0^\varepsilon d\varepsilon' D_t(\varepsilon')\right] \frac{d}{d\varepsilon} \log \left(1 \mp e^{-\beta(\varepsilon-\mu)}\right). \quad (21.46)$$

The first term vanishes (you must check this$^{257*}$), so ((21.44) will be used)

$$PV = \pm k_B T \int d\varepsilon \left[\int_0^\varepsilon d\varepsilon' D_t(\varepsilon')\right] \frac{d}{d\varepsilon} \log \left(1 \mp e^{-\beta(\varepsilon-\mu)}\right) \quad (21.47)$$

$$= \pm \frac{2}{3} k_B T \int d\varepsilon \varepsilon D_t(\varepsilon) \frac{d}{d\varepsilon} \log \left(1 \mp e^{-\beta(\varepsilon-\mu)}\right) \quad (21.48)$$

$$= \pm \frac{2}{3} k_B T \int d\varepsilon \varepsilon D_t(\varepsilon) \frac{\pm \beta e^{-\beta(\varepsilon-\mu)}}{1 \mp e^{-\beta(\varepsilon-\mu)}} \quad (21.49)$$

$$= \frac{2}{3} \int d\varepsilon D_t(\varepsilon) \frac{\varepsilon}{e^{\beta(\varepsilon-\mu)+1}} = \frac{2}{3} E. \quad (21.50)$$

Now, a bomb making question. There is an isolated metal container of volume $V$ and energy $E$ with $N$ fermions whose chemical potential is $\mu = 10$ eV. These fermions are adiabatically converted into bosons. Can we use this mechanism to make a bomb? Notice that since $E$ and $V$ are unchanged, the pressure does not change. [Hint: You must know 1 eV is roughly equivalent to $10^5$ K.]

$^{257*}$ $\varepsilon \to 0$ for the boson case is only slightly tricky, but $\varepsilon^{1/2} \log \varepsilon \to 0$ saves the day.
Q21-1. In a big box is an ideal gas A whose mass is $m$ per molecule. Inside the box is a surface with $N$ absorption sites that can accommodate at most one A molecules per site. When one A is absorbed, its energy is $-\varepsilon$ ($\varepsilon > 0$, i.e., $\varepsilon$ lower than the free state) and has $z$ internal states with the same energy.

When the pressure of the box is $P$, and the temperature is $T$, what is the fraction $\theta$ of the surface occupied by A? Or, more concretely, find the factor $X$ in the following formula

$$\theta = \frac{P}{P + P_Q/X}, \quad (21.51)$$

where $P_Q = n_Qk_B T$ with $n_Q = (2\pi mk_B T/h^2)^{3/2}$. The formula (21.51) is called the *Langmuir isotherm*.

**Soln.**

(1) We regard the gas phase as a chemical reservoir, whose chemical potential $\mu$ may be calculated (or found in the lecture notes) later. Under this chemical potential the grand canonical partition function of the absorbing surface can be written as

$$\Xi = \prod_{i=1}^{N} \Xi_i, \quad (21.52)$$

where $\Xi_i$ is the ‘grand canonical partition function for the $i$th absorption center.’ We may write

$$\Xi_i = 1 + ze^{-\beta(-\varepsilon-\mu)} = 1 + ze^{\beta(\varepsilon+\mu)} . \quad (21.53)$$

The expectation value of the total number $M$ of the absorbed A molecules is $N$ times expected number of A at a single absorption center:

$$M = \frac{\partial\log \Xi}{\partial \beta \mu} = N \frac{ze^{\beta(\varepsilon+\mu)}}{1 + ze^{\beta(\varepsilon+\mu)}}. \quad (21.54)$$

Now, let us obtain (or copy) $\mu$:

$$\mu = k_B T \log \frac{P}{k_B T n_Q}, \quad (21.55)$$

or

$$e^{\beta \mu} = \frac{P}{k_B T n_Q} \equiv \frac{P}{P_Q}, \quad (21.56)$$

where $n_Q = (2\pi mk_B T/h^2)^{3/2}$. Therefore, $\theta = M/N$ reads

$$\theta = \frac{z\alpha(P/P_Q)}{1 + z\alpha(P/P_Q)} = \frac{P}{P + P_Q/z\alpha}. \quad (21.57)$$
where \( \alpha = e^{\beta \varepsilon} \). That is,
\[
X = z e^{\beta \varepsilon}.
\]  

(21.58)

**Q21-2.** There is a system in which each particle can assume only three states with energies 0, \( \varepsilon \) and \( \varepsilon \) (\( \varepsilon > 0 \), i.e., excited states are degenerate). There are two identical particles without spin.

(1F) When the particles are fermions, write down the canonical partition function (I recommend you to make a table of all the microstates).

(2F) Find the probability of finding \( N (= 0, 1, 2) \) particles in the ground state.

(3F) Compute the average occupation number \( N_0 \) of the ground state. Are the limits \( T \to \infty \) and \( T \to 0 \) reasonable?

(1-3B) Repeat the same problems assuming that the particles are bosons.

(4) In the high temperature limit what is the most important observation?

**Soln.**

Here ‘degenerate’ means that the energies happen to be identical but the states are clearly distinguishable like the three \( 2p \) orbits in the hydrogen atom.

(1F) To compute the canonical partition function, you must itemize all the microstates.

<table>
<thead>
<tr>
<th>microstate</th>
<th>0</th>
<th>( \varepsilon )</th>
<th>( \varepsilon )</th>
<th>total energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>( 2 \varepsilon )</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>( \varepsilon )</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>( 2 \varepsilon )</td>
</tr>
</tbody>
</table>

Hence,

\[
Z = 2 e^{-\beta \varepsilon} + e^{-2 \beta \varepsilon}.
\]

(2F) Let us write the desired probabilities as \( P(N) \).

\[
P(1) = \frac{2 e^{-\beta \varepsilon}}{2 e^{-\beta \varepsilon} + e^{-2 \beta \varepsilon}} = \frac{2}{2 + e^{-\beta \varepsilon}}, \quad P(0) = \frac{e^{-2 \beta \varepsilon}}{2 e^{-\beta \varepsilon} + e^{-2 \beta \varepsilon}} = \frac{1}{e^{\beta \varepsilon} + 2}, \quad P(2) = 0.
\]

(3F)

\[
\langle N_0 \rangle = P(1) = \frac{2}{2 + e^{-\beta \varepsilon}}
\]

\( T \to \infty \): \( N_0 = 2/3 \) (yes, all the states are equally probable).

\( T \to 0 \): \( N_0 = 1 \) (yes, the lowest level must surely be occupied).

(1B) To compute the canonical partition function, you must itemize all the mi-
Hence,

\[ Z = 1 + 2e^{-\beta \varepsilon} + 3e^{-2\beta \varepsilon}. \]

(2B)

\[ P(1) = \frac{2}{Z}e^{-\beta \varepsilon}, \quad P(0) = \frac{3}{Z}e^{-2\beta \varepsilon}, \quad P(2) = \frac{1}{Z}. \]

(3B)

\[ N_0 = \frac{2 + 2e^{-\beta \varepsilon}}{1 + 2e^{-\beta \varepsilon} + 3e^{-2\beta \varepsilon}}. \]

\( T \to \infty: \) \( N_0 = \frac{2}{3} \) (yes, all the states are equally probable, and must be the same as the fermion case).

\( T \to 0: \) \( N_0 = 2 \) (yes, all the particles must be there).

(4) Both agree as noted above.

**Q21-3.** There are 120 identical spinless bosons whose \( n \)-th one-particle state has an energy \( \varepsilon_n = n\varepsilon \) \( (n \in \mathbb{N}) \). These particles do not interact. When the system is in equilibrium with the particle reservoir (chemostat) of temperature \( T \) and chemical potential \( \mu \), on the average 119 particles occupy the one-particle ground state \( (n = 0) \), and one particle occupies the one-particle first excited state \( (n = 1) \). The other one-particle states are negligibly occupied.

(1) Find the chemical potential \( \mu \) in terms of \( \varepsilon \) (or compute \( \mu/\varepsilon \)).

(2) Is the second excited state occupied only negligibly? Compute \( \langle n_2 \rangle \). \( \langle n_2 \rangle / \langle n_1 \rangle \) is not terribly small, so you might think that the problem is not self-consistent. Give your comment on this observation.

**Soln.**

(1) Since

\[ \langle n_0 \rangle = \frac{1}{e^{-\beta \mu} - 1} = 119, \quad \text{(21.59)} \]

\[ \langle n_1 \rangle = \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} = 1, \quad \text{(21.60)} \]
we have

\[-\beta \mu = \log(120/119) = 0.00836825, \quad (21.61)\]
\[\beta(\varepsilon - \mu) = \log 2 = 0.693147. \quad (21.62)\]

Hence, \(\beta = 0.684779/\varepsilon\) and \(\mu = -0.0122204\varepsilon\). Clearly recognize that \(\mu\) is negative (does not exceed the ground state energy)!

(2) \(\beta(2\varepsilon - \mu) = 2\beta(\varepsilon - \mu) - (-\beta \mu) = 2 \times 0.693147 - 0.00836825 = 1.37793\), so \(\langle n_2 \rangle = 0.337\). Actually, \(\langle N \rangle\) is about 0.5 articles more than 120. This is inevitable because \(\log 120/120 = 0.04\), so a few % error is actually expected. Our result is much better than this crude error estimate.

**Q21-4.** [Ensemble review]
For a fluid the Gibbs free energy \(G = E - ST + PV\).
(1) Find the generalized canonical partition function \(Q\) that gives \(G\) directly as

\[G = -k_B T \log Q. \quad (21.63)\]

Needless to say, you must assume \(T\) and \(P\) are given.
(2) Compute \(Q\) for a classical ideal gas and obtain the ideal gas equation of state. You may use the canonical partition function \(Z(T, V)\) for the ideal gas already computed. (The needed summation over \(V\) (volume) is actually an integration.)

**Soln.**
(1) Since

\[G = E - TS + PV, \quad (21.64)\]

so

\[Q = \sum_{E,V} w(E,V)e^{-\beta(E+PV)} = \sum_V Z(T,V)e^{-\beta PV}, \quad (21.65)\]

where \(Z\) is the usual canonical partition function

\[Z = \frac{1}{N!} \left(\frac{2\pi m k_B T}{\hbar^2}\right)^{3N/2} V^N. \quad (21.66)\]

This sum is actually an integration.
(2) For the \(N\) particle case, let us write \(Q\) with suffix \(N\) as \(Q_N\):

\[Q_N = \frac{1}{N!} \left(\frac{2\pi m k_B T}{\hbar^2}\right)^{3N/2} \int_0^\infty dV V^N e^{-\beta PV} \quad (21.67)\]
The easiest way to compute this is to consider (if you know the Γ-function, it is easy)

$$
\sum_N \int_0^\infty dV \frac{1}{N!} V^N \alpha^N e^{-\beta PV} = \int dV e^{-(\beta P-\alpha)V} = \frac{1}{\beta P - \alpha} = \frac{1}{\beta P} \sum_N \left( \frac{\alpha}{\beta P} \right)^N,
$$

(21.68)

that is, to use the generating function. Hence,

$$
\int_0^\infty dV \frac{1}{N!} V^N e^{-\beta PV} = \left( \frac{k_B T}{P} \right)^{N+1}.
$$

(21.69)

Therefore,

$$
Q_N = \left( \frac{2\pi m k_B T}{\hbar^2} \right)^{3N/2} \left( \frac{k_B T}{P} \right)^{N+1}.
$$

(21.70)

We need \(dG = -SdT + VdP\). From

$$
G = -k_B T \log Q_N = (N+1)k_B T \log P + \cdots.
$$

(21.71)

we get

$$
V = \left. \frac{\partial G}{\partial P} \right|_T = k_B T (N+1) \frac{1}{P}
$$

(21.72)

You might complain that \(N+1\) appears instead of \(N\), but the equivalence of ensembles ignores \(\log N\) against \(N\), so ignoring 1 against \(N\) is much less sinful.

**Q21-5.** If we compare the pressure \(P_{BE}\) of an ideal boson gas and that \(P_{MB}\) of an ideal classical gas under the same \(V, T, N\), \(P_{MB} > P_{BE}\). Mimicking the Fermi-Dirac case (i.e., \(P_{FD} > P_{MB}\)) explained in the lecture notes, demonstrate this inequality.

**Soln.**

\(\beta P_{MB} V = N\), so we wish to compare \(\beta P_{BE} V\) and \(N\) (since \(\mu\)'s are not the same for the two cases, we cannot immediately comparer the pressure formulas).

$$
\beta P_{BE} V = -\sum_i \log(1 - e^{-\beta(\epsilon_i - \mu)}),
$$

(21.73)

and

$$
N = \sum_i \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1} = \sum_i \frac{e^{-\beta(\epsilon_i - \mu)}}{1 - e^{-\beta(\epsilon_i - \mu)}}.
$$

(21.74)

Here, you must use consistently the formula for bosons.

Let us compare the corresponding summand terms in \(P\) and \(N\), writing \(x = e^{-\beta(\epsilon_i - \mu)}\). Let

$$
f(x) = \frac{x}{1 - x} + \log(1 - x).
$$

(21.75)
\[ f(0) = 0. \]

\[ f'(x) = \frac{1}{1 - x} + \frac{x}{(1 - x)^2} - \frac{1}{1 - x} = \frac{x}{(1 - x)^2}, \]  \hspace{1cm} (21.76)

which is positive for \( x > 0 \). Therefore, \( f(x) > 0 \), if \( x > 0 \). QED.

Q21-6. There is a mixture of \( \frac{N}{3} \) non-interacting fermions and \( 2\frac{N}{3} \) non-interacting bosons in a container of volume \( V \). The total internal energy is \( E \). What is the total pressure \( P \) of the mixture? [Hint: Dalton’s law of partial pressures applies. You must clearly state your logic to support your answer.]

Soln.

Let \( P_F \) and \( P_B \) be the respective partial pressures, and \( E_F \) and \( E_B \) internal energies of the respective components. Then,

\[ P_F V = \frac{2}{3} E_F, \quad P_B V = \frac{2}{3} E_B, \]

so

\[ (P_F + P_B)V = \frac{2}{3}(E_F + E_B) \Rightarrow PV = \frac{2}{3} E, \]

which is due to additivity of internal energy and the law of partial pressures. Hence, \( P = 2E/3V \).
22 Ideal quantum gases at very low temperatures

Summary
* Elementary low temperature behaviors of non-interacting particle systems are discussed.
* We will guess low temperature behaviors of $E, S, \mu$ for free fermions.

Key words
Fermi energy, Bose-Einstein condensation, condensate

What you should be able to do
* You must know the general value of $\varepsilon_F$.
* You should be able to calculate various quantities for $T = 0$ fermion.
* For $E, \mu T \neq 0$ corrections start with the terms of order $T^2$. You must be able to explain why.
* Understand why $C_V \propto T$ for fermions close to $T = 0$.
* Remember the shape and rough scales of the derivative of the Fermi-Dirac distribution.
* Understand why Bose-Einstein condensation occurs.

22.1 Noninteracting fermion pressure at $T = 0$
The equation of state reads $PV = 2E/3$, so let us compute the internal energy at $T = 0$. The one particle states are completely filled up to the chemical potential $\mu(0)$ at $T = 0$, which is determined by

$$N = \int_0^{\mu(0)} d\varepsilon D_{\xi}(\varepsilon) = \gamma_0 V \int_0^{\mu(0)} \varepsilon^{1/2} d\varepsilon = \frac{2}{3} \gamma V \mu(0)^{3/2}. \quad (22.1)$$

Here, $D_{\xi}(\varepsilon) = \gamma_0 V \varepsilon^{1/2}$ with $\gamma_0 = 2\pi(2m/\hbar^2)^{3/2}$. We do not need details, but

$$\mu(0) \propto n^{2/3} \quad (22.2)$$

is worth remembering, where $n$ is the number density.

The internal energy at $T = 0$ is given by

$$E(0) = \int_0^{\mu(0)} d\varepsilon D_{\xi}(\varepsilon) \varepsilon = \gamma_0 V \int_0^{\mu(0)} \varepsilon^{3/2} d\varepsilon = \frac{2}{5} \gamma_0 V \mu(0)^{5/2} = \frac{3}{5} N \mu(0). \quad (22.3)$$
The last formula should be obtainable by dimensional analysis except for the numerical factor. This implies

$$PV = \frac{2}{5} N \mu(0).$$

(22.4)

Notice that this is usually very large. You might have realized that at $T = 0$, $E = -PV + \mu(0)N$, so our familiar $E = 2PV/3$ gives us everything we wish.

### 22.2 Low temperature specific heat of electrons (fermions)

Let us intuitively discuss the electronic heat capacity of metals at low temperatures. We may assume that the Fermi-Dirac distribution is (almost) a step function. We can infer from the width of the ‘avalanche region’ of the cliff of the Fermi-Dirac distribution that the number of excitable electrons is $\sim N k_B T$ at the temperature around $T$. We know generally that the specific heat is proportional to the number of the excitable degrees of freedom, so $C_v \propto T$ at lower temperatures. Thus, at sufficiently low temperatures this dominates the heat capacity of metals (where $T^3$ coming from the lattice vibration is much less than $T$).

![Diagram](image)

Figure 22.1: $C_v$ is proportional to the number of degrees of freedom that may be excited at temperature $T$. For photons and phonons all the particles occupying the one particle states up to the energy $\sim k_B T$ can be excited, so in 3D the specific heat is proportional to $T^3$ (this is also true for superrelativistic bosons). For fermion systems, among the occupied one particle states, only the particles within the width $\sim k_B T$ near the top of the occupied states can be excited, so $C_v \propto T$. These ideas may be used in any dimensional space.
22.3 Low temperature approximation for fermions

To get this result more quantitatively, we need a way to estimate the contribution of the cliff width. Let us look at the formula for $N$:

$$N = \int_{0}^{\infty} D(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)} + 1}. \quad (22.5)$$

Figure 22.2: The derivative of the Fermi distribution. Its width is about $5k_B T$ and the height is $\beta/4$.

Let $f(\varepsilon) = 1/(e^{\beta(\varepsilon-\mu)} + 1)$. Since we know $-f'(\varepsilon)$ is concentrated sharply around $\varepsilon = \mu$ (Fig. 22.2), we wish to exploit this fact:

$$\int_{0}^{\infty} \left[ \int_{0}^{\varepsilon} D(\varepsilon') d\varepsilon' \right] f(\varepsilon) d\varepsilon = \int_{0}^{\varepsilon} D(\varepsilon') d\varepsilon' f(\varepsilon) \bigg|_{\varepsilon=0} - \int_{0}^{\infty} d\varepsilon \left[ \int_{0}^{\varepsilon} D(\varepsilon') d\varepsilon' \right] f'(\varepsilon)$$

$$= -\int_{0}^{\infty} d\varepsilon \left[ \int_{0}^{\varepsilon} D(\varepsilon') d\varepsilon' \right] f'(\varepsilon). \quad (22.6)$$

Now, $f'$ is localized around $\mu$, so we need the quantity in $[ ]$ only near $\varepsilon = \mu$. Therefore, let us Taylor-expand it as follows:

$$-\int_{0}^{\infty} \left[ \int_{0}^{\mu} D(\varepsilon') d\varepsilon' + D(\mu)(\varepsilon - \mu) + \frac{1}{2} D'(\mu)(\varepsilon - \mu)^2 + \cdots \right] f'(\varepsilon) d\varepsilon$$

$$= \int_{0}^{\mu} D(\varepsilon') d\varepsilon' - \frac{1}{2} D'(\mu) \int_{0}^{\infty} (\varepsilon - \mu)^2 f'(\varepsilon) d\varepsilon + \cdots. \quad (22.7)$$

Detailed calculation is not given here, but it is clear that the correction is of order $T^2$; the integral in the second term has the dimension of energy squared, so it must be proportional to $(k_B T)^2$. If we wish to compute $E$ in the above calculation, $D$ is replaced by $\varepsilon D$, but the expansion method is exactly the same, so the correction
term is proportional to $T^2$. That is, although we do not go through any detailed calculation, we can conclude with a certain positive number $\alpha$ (because $E$ must increase with $T$) that

$$E(T) = E(0) + \frac{1}{2}\alpha T^2 + \cdots.$$  \hspace{1cm} (22.8)

Therefore, as we have expected above, $C_V = \alpha T$ for sufficiently small $T$.

### 22.4 Low temperature entropy of fermion systems
We know

$$\left. \frac{\partial S}{\partial T} \right|_V = \frac{C_V}{T} = \alpha.$$  \hspace{1cm} (22.9)

This implies that under $V, N$ constant condition $S(T) = S(0) + \alpha T$, but $S(0) = 0$ is assumed usually, so we conclude that for sufficiently small $T$

$$S(T) = C_V.$$  \hspace{1cm} (22.10)

### 22.5 Low temperature behavior of chemical potential
Now, let us study the $T$ dependence of $\mu$. You may probably guess that $\mu(T) = \mu(0) - O[T^2]$ from the above calculation. We know $\mu = G/N = (E - ST + PV)/N = (5E/3 - ST)/N$ for noninteracting systems. Therefore, we confirm our guess:

$$N \mu = \frac{5}{3} \left( E(0) + \frac{1}{2}\alpha T^2 \right) - \alpha T^2 = \frac{5}{3} E(0) - \frac{1}{6} \alpha T^2.$$  \hspace{1cm} (22.11)

What happens if the spatial dimension is 1? It is an increasing function of $T$ for sufficiently low $T$.

### 22.6 Difficulty of continuum approximation at low temperatures for bosons
Next, let us study the free boson system. Let us take the ground state energy of the system to be the origin of energy. Then, the chemical potential cannot be positive. The total number of particles in the system of free bosons is given by

$$N = \sum_i \frac{1}{e^{\beta(\epsilon_i - \mu)}} - 1.$$  \hspace{1cm} (22.12)
If $T$ is sufficiently small, the first term corresponding to the single-particle ground state can become very large (see Fig. 22.3), so in general it is dangerous to approximate (22.12) by integral with the aid of a smooth density of state as in the fermion case (in case of fermions, each term cannot be larger than 1, so there is no problem at all in this approximation).

Let us look at the difficulty in approximating (22.12) by an integration in 3d. In this case the density of states has the form $D_t(\varepsilon) = \gamma V \varepsilon^{1/2}$ with some positive constant $\gamma$. Let us write the continuous approximation to (22.12) as $N_1$:

$$N_1(T, \mu) \equiv \gamma V \int_{0}^{\infty} d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-\mu)} - 1}. \quad (22.13)$$

$N_1$ is a function of $T$ and $\mu$. It is an increasing function of $T$, and also an increasing function of $\mu$. For a given $T$, if we can choose $\mu$ (which must be negative) satisfying $N = N_1$, then we can describe the system with a continuous approximation of the grand canonical ensemble with this $\mu$.

Now, let us decrease the temperature. Then, $N_1$ decreases, so to keep $N_1 = N$ we must increase $\mu$. However, we cannot indefinitely increase $\mu$; $\mu = 0$ is the upper limit. Is there any guarantee that before reaching this limit, $N = N_1(T, \mu)$ may always be satisfied?

Since $N_1(T, \mu) \leq N_1(T, 0)$,

$$\gamma V \int_{0}^{\infty} d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-\mu)} - 1} \leq \gamma V (k_B T)^{3/2} \int_{0}^{\infty} dz \frac{z^{1/2}}{e^z - 1}. \quad (22.14)$$

The integral on the right-hand side is finite. That is, with a positive constant $A$ we may write

$$N_1(T, \mu) \leq A V T^{3/2}. \quad (22.15)$$

The equality holds when $\mu = 0$. $N_1$ can be made indefinitely close to 0 by reducing $T$. However, the system should have $N$ bosons independent of $T$, so there must be a temperature $T_c$ at which

$$N = N_1(T_c, 0) \quad (22.16)$$

and for $T < T_c$

$$N > N_1(T, 0). \quad (22.17)$$

This is the difficulty of continuum approximation for low temperature bosons in 3-space.

### 22.7 Bose-Einstein condensation

The temperature $T_c$ is called the *Bose-Einstein condensation temperature*; below this
the continuous approximation breaks down. Notice that \( T_c \) is determined by (22.16) and

\[
T_c \propto n^{2/3}.
\]  

(22.18)

Since the system must have \( N \) particles, the remaining \( N_0 = N - N_1 \) must occupy some one particle state. The only possibility is the ground state which is not properly taken into account by the continuous approximation.

![Figure 22.3: If \( T < T_c \), where \( T_c \) is the Bose-Einstein condensation temperature, then the ground state is occupied by \( N_0 = O[N] \) particles, so the approximation of (22.12) by an integral becomes grossly incorrect.

A macroscopic number \( N_0 (= N - N_1) \) of particles fall into the lowest energy one particle state (see Fig. 22.3). This phenomenon is called a Bose-Einstein condensation. Notice that except for the ground state no other one particle states are occupied by macroscopic numbers of particles under any condition. Only the one-particle ground state can be occupied by a macroscopic number of particles below \( T_c \). Here, ‘macroscopic’ implies that \( N_0/N \) is a positive number in the large system size limit (\( N \to \infty \) limit, the thermodynamic limit).

### 22.8 Non-condensate population

From (22.14), we know that \( N_1 \) is an increasing function of \( \mu \), but we cannot increase \( \mu \) indefinitely; \( \mu \) must be non-positive. Hence, at or below a certain particular temperature \( T_c \) \( \mu \) vanishes. That is, at \( T = T_c \) the equality must hold in (22.14), so \( T_c \) is fixed by the condition (22.16), that is,

\[
N = C(k_B T_c)^{3/2} \int_0^\infty dz \frac{z^{1/2}}{e^{z} - 1}.
\]  

(22.19)
Hence, we get for $T \leq T_c$ (Fig. 22.4)

$$N_1 = N \left( \frac{T}{T_c} \right)^{3/2}. \tag{22.20}$$

Thus the condensate population reads for $T \leq T_c$.

$$N_0 = N \left[ 1 - \left( \frac{T}{T_c} \right)^{3/2} \right]. \tag{22.21}$$

![Figure 22.4: The ratio $N_1/N$ of non-condensate atoms has a singularity at the Bose-Einstein condensation point $T_c$. The lower panel describes the chemical potential.]

**22.9 Bose-Einstein condensation does not occur in 2 and 1 dimensions**

No Bose-Einstein condensation occurs in one and two dimensional free spaces, because $N_1$ is not bounded from above. For example, in 2D

$$N_1 = \int_0^\infty D_2(\varepsilon) \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} d\varepsilon, \tag{22.22}$$

where $D_2(\varepsilon)$ is the density of states of a single particle in 2-space. Let us repeat our quick derivation:

$$\int_0^\varepsilon d\varepsilon D_2(\varepsilon) = \frac{V}{\hbar^2} \int_{p^2/2m \leq \varepsilon} dp = \frac{2\pi V}{\hbar^2} \int_0^{\sqrt{2m\varepsilon}} p dp, \tag{22.23}$$

so

$$D_2(\varepsilon) = \frac{2\pi V}{\hbar^2} \sqrt{2m\varepsilon} \frac{d\sqrt{2m\varepsilon}}{d\varepsilon} = cV, \tag{22.24}$$
where \( c \) is a constant. Therefore,

\[
N_1 \propto V \int_0^\infty \frac{1}{e^{(\varepsilon - \mu)} - 1} d\varepsilon. \tag{22.25}
\]

We know \( N_1 \) must be an increasing function of \( \mu \) and the largest possible \( \mu \) is zero for bosons, so

\[
\int_0^\infty \frac{1}{e^{(\varepsilon - \mu)} - 1} d\varepsilon \leq \int_0^\infty \frac{1}{e^{\beta \varepsilon} - 1} d\varepsilon. \tag{22.26}
\]

The integral on the right-hand side blows up from the contribution close to \( \varepsilon = 0 \); there \( 1/(e^{\beta \varepsilon} - 1) \simeq 1/\beta \varepsilon \), so the integral diverges logarithmically. Therefore, for any \( N \) and \( T \), we can find \( \mu < 0 \) such that \( N = N_1 \). Thus, there is no Bose-Einstein condensation.

Why don’t you check the 1D case?

### 22.10 Continuum approximation is always valid for \( E \) and \( P \)

Notice that the integral expression for \( E \) or \( PV \) is still all right, because for these quantities the ground state does not contribute at all.\(^{258}\)

### 22.11 Low temperature heat capacity of bose systems

The Bose-Einstein condensate (i.e., \( N_0 \)) does not contribute to internal energy, so we may use the continuum approximation to compute the internal energy. Below \( T_c \) we may set \( \mu = 0 \), so

\[
E = \int_0^\infty d\varepsilon D(\varepsilon) \frac{\varepsilon}{e^{\beta \varepsilon} - 1}. \tag{22.27}
\]

In 3-space, we know \( D(\varepsilon) = \gamma V \varepsilon^{1/2} \) with \( \gamma \) being a positive constant. Therefore, for \( T < T_c \)

\[
E = \gamma V \int_0^\infty d\varepsilon \varepsilon^{1/2} \frac{\varepsilon}{e^{\beta \varepsilon} - 1} = \gamma V \beta^{-5/2} \int_0^\infty d(\beta \varepsilon) \frac{(\beta \varepsilon)^{3/2}}{e^{\beta \varepsilon} - 1} \propto VT^{5/2}. \tag{22.28}
\]

(More easily, we can simply count the power of \( \varepsilon \). Here, we have \( d\varepsilon, \varepsilon^{1/2} \) and \( \varepsilon \), so \( \varepsilon^{5/2} \) is the ‘dimension of the integral.’ The only relevant quantity with the dimension

\(^{258}\)Accurately speaking, a careful calculation shows that the ground-state contributions to \( E \) and \( P \) are of order \( \log N \), which we may ignore. If not, the grand canonical formalism cannot be applied in any case.
of energy is $k_B T$, so this integral must be proportional to $T^{5/2}$.) From this the low temperature heat capacity is

$$C_V \propto \left( \frac{T}{T_c} \right)^{3/2}.$$ 

(22.29)

This goes to zero with $T$ as required by the third law.

Notice that $C_V$ is proportional to the number of degrees of freedom excitable at around $T$. 

326
Q22-1. Find the density of states $D(\varepsilon)$ (for the translational degrees of freedom) of a single particle in the volume $V$ in 2-dimensional space with the super-relativistic dispersion relation $\varepsilon = c|p|$.

**Soln.**
Here, I give the most general solution with (fairly detailed) explanation: in $d$-space with the dispersion relation $\varepsilon = \alpha |p|^{\gamma}$, where $\alpha$ is a positive constant. Our strategy is always the same. If we can study the classical phase volume, dividing it with $h^d$ in $d$-space, we can obtain the number of states for a single particle. The single-particle states with energy not exceeding $\varepsilon$ corresponds to those with the momenta satisfying $|p| \leq (\varepsilon/\alpha)^{1/\gamma}$. Therefore, the number of the single-particle states with energy less than $\varepsilon$ may be written in two ways:

$$\int_0^\varepsilon d\varepsilon \, D(\varepsilon) = \frac{1}{h^d} \int d^d q \int_{|p| \leq (\varepsilon/\alpha)^{1/\gamma}} d^d p.$$

The right-hand side may be rewritten by computing the position-coordinate integral (that gives $V$) and by introducing the polar coordinate system

$$\int_0^\varepsilon d\varepsilon \, D(\varepsilon) = \frac{1}{h^d} V \int_0^{(\varepsilon/\alpha)^{1/\gamma}} S_{d-1} p^{d-1} dp,$$

where $S_{d-1}$ is the volume of $d - 1$-unit sphere, whose general form may be found in my graduate course notes. To obtain $D(\varepsilon)$, we simply differentiate the above identity:

$$D(\varepsilon) = \frac{V}{h^d} S_{d-1} \frac{\varepsilon^{(d-1)/\gamma} d(\varepsilon/\alpha)^{1/\gamma}}{d\varepsilon} = \frac{V}{h^d} S_{d-1} \frac{\varepsilon^{(d-1)/\gamma} d\varepsilon^{1/\gamma}}{d\varepsilon} = \frac{S_{d-1}}{\gamma h^d} \frac{V \varepsilon^{d/\gamma - 1}}{\alpha^{d/\gamma}}.$$

For $d = 3$ and $\gamma = 2$ (the usual particle in 3-space), $S_2 = 4\pi$, $\alpha = 1/2m$ and we recover the formula we know well

$$D(\varepsilon) = 2\pi \frac{V}{h^3} (2m)^{3/2} \varepsilon^{1/2}.$$

For the problem case $d = 2$, $\gamma = 1$, $S_1 = 2\pi$, and $\alpha = c$, so

$$D(\varepsilon) = 2\pi \frac{V \varepsilon}{h^2 c^2}.$$

Of course, this has ‘almost’ been given when we studied Planck’s formula.

---

\(^{259}\text{f}(x) = g(x)\) means \(f'(x) = g'(x)\) if the first equality is an identity (and differentiable) in \(x\).
Q22-2. There is a 2-dimensional harmonic trap $U = \frac{1}{2} \alpha x^2$ where $x$ is the distance from the origin, and $\alpha$ is a positive constant. We know the single particle energy levels in this trap are denoted as

$$\varepsilon = \hbar \omega (1 + n_1 + n_2),$$  \hspace{1cm} (22.30)

where $n_1, n_2 \in \mathbb{N} = \{0, 1, 2, \cdots\}$, and $\omega$ is a positive constant.

1. The density of states $D(\varepsilon)$ is the number of states with energy between $\varepsilon$ and $\varepsilon + d\varepsilon$. Here, however, to make the one particle ground state to be with zero energy, let us subtract the zero-point energy in the following. Therefore, we know

$$\int_{0}^{\varepsilon} d\varepsilon' D(\varepsilon') = \sum_{n_1 + n_2 \in [0, \varepsilon / \hbar \omega]} 1. \hspace{1cm} (22.31)$$

Noting that the sum on the right-hand side is essentially the area of the shaded triangle in the following figure, obtain $D(\varepsilon)$.

![Figure 22.5: The relation between $n_1$, $n_2$ and $\varepsilon/\hbar \omega$. Here, the zero-point energy ($\hbar \omega$) has been subtracted from $\varepsilon$.](image)

(2) Is there a Bose-Einstein condensation in this 2D trap? [Hint. Compute $N_1$ and study whether it is finite or not for $\mu = 0$. Mimic our argument in 3D free space.]

**Soln.**

1. As explained, the right-hand side must be the area of the triangle, so

$$\int_{0}^{\varepsilon} d\varepsilon' D(\varepsilon') = \frac{1}{2} \left( \frac{\varepsilon}{\hbar \omega} \right)^2.$$

Differentiating this with $\varepsilon$, we obtain

$$D(\varepsilon) = \frac{\varepsilon}{(\hbar \omega)^2} \propto \varepsilon.$$  \hspace{1cm} (22.32)

$^{260}$This can be realized on graphene.
(2) Warning: $\mu = 0$ below $T_c$ is when we choose the ground state energy to be zero. We have already removed the zero-point energy, so we can consider $\mu = 0$.

Let us compute (the expectation value of) the total number $N_1$ of particles in the excited states for $\mu = 0$:

$$N_1 = \int_0^\infty d\varepsilon \frac{1}{e^{\beta\varepsilon}} \left[ \int_0^\infty d\varepsilon \left( \frac{\varepsilon}{e^{\beta\varepsilon}} - 1 \right) = k_B T \int_0^\infty dz \frac{z}{e^z - 1} \right].$$

This integral is finite: the dangerous contribution comes from small $z$, because for large $z$ the integrand is exponentially small. For small $z$ the integrand tends to a constant, so the integral is finite, and $N_1$ can be indefinitely small for sufficiently small $T$. Therefore, $N_0 = N - N_1$, the condensate population, must be macroscopic. That is, we can expect a Bose-Einstein condensation.

**Q22-3.** There are $2N$ non-interacting fermions in a container of volume $V$. While the volume is isoenergetically halved (that is $V \rightarrow V/2$ while $E$ is kept constant), two fermions react to make a single boson. After the reaction completed, all the fermions are converted into non-interacting bosons (i.e., $N$ bosons in volume $V/2$ with internal energy $E$). Assume that the chemical potential of the initial fermion system is $0.7$ eV and its temperature is $T = 0$. Is the final temperature higher than $5 \times 10^3$ K? You MUST justify your answer, since guessing the answer may not be very hard. [Hint: the ‘bomb question.’]

**Soln.**
The total energy of the original system would be estimated from the $T = 0$ formula as $E = (3/5)2N\mu(0)$. After the reaction we have $N$ bosons. We know for given $T$, $N$ and $V$, $P_{MB} > P_{BE}$, but if $E$, $N$ and $V$ are the same $P_{MB} = P_{BE} = (2/3)E/V$. From $P_{MB}(V/2) = Nk_B T_{MB} = (2/3)E$, we have

$$T_{MB} = \frac{2E}{3Nk_B} = \frac{2}{3} \frac{6N\mu}{5Nk_B} = \frac{4\mu}{5k_B}.$$

This is equal to $0.56$ eV/$k_B \cong 6000$ K $> 5 \times 10^3$ K. That is, classical gas would have about this temperature.

To estimate the boson system temperature is actually rather delicate, but the following argument tells us that the final temperature is definitely higher than $5 \times 10^3$ K. The pressure of any gas is an increasing function of $T$. If the temperature is the same, that is, if $T$, $N$, and $V$ are the same, $P_{MB} > P_{BE}$, but in our case they are identical, so the actual temperature of the final bosonic system must be higher than $T_{MB}$.

**Q22-4.** [The following questions we may have already discussed in a lecture. Still give justifying arguments (or formulas) to your answers (write your own argument).]
Watch out for trivial questions.

Assume that the particles do not interact, and answer the following questions for both ideal bosons and ideal fermions (both without any internal degree of freedom).

(1) The volume $V$ is increased under constant internal energy. Does the temperature decrease? Assume that the initial temperature is sufficiently low (below $T_c$ for bosons).

(2) The volume $V$ is increased under constant entropy. Does the temperature decrease?

(3) The volume $V$ is increased under constant temperature. Does the pressure decrease?

**Soln.**

(1) If the volume is increased, the level spacings decrease.

(F) For fermions, if this happens at a very low temperature, then particles must be excited to go beyond the Fermi energy. Thus, $T$ increases. More quantitatively, let us ‘head-on’ consider

$$\frac{\partial T}{\partial V} \bigg|_E = \frac{\partial(T, E)}{\partial(V, E)} = \frac{\partial(T, E)}{\partial(V, T)} \frac{\partial(V, T)}{\partial(V, E)} = - \frac{1}{C_V} \frac{\partial E}{\partial V} \bigg|_T. \tag{22.32}$$

Using the Gibbs relation, we get

$$\frac{\partial E}{\partial V} \bigg|_T = T \frac{\partial S}{\partial V} \bigg|_T - P, \tag{22.33}$$

but

$$\frac{\partial S}{\partial V} \bigg|_T = \frac{\partial(S, T)}{\partial(V, T)} = \frac{\partial(S, T)}{\partial(V, P)} \frac{\partial(V, P)}{\partial(V, T)} = \frac{\partial P}{\partial T} \bigg|_V \tag{22.34}$$

with the aid of a Maxwell’s relation. Using $P = 2E/3V$,

$$\frac{\partial S}{\partial V} \bigg|_T = \frac{2C_V}{3V}. \tag{22.35}$$

Thus,

$$\frac{\partial E}{\partial V} \bigg|_T = \frac{2}{3V} (T C_V - E). \tag{22.36}$$

For free fermions, we know $E$ at low temperatures has a big $T$-independent chunk $E_0$, so for sufficiently low temperatures, this derivative must be negative. Hence, (22.32) must be positive.

(B) We know classically $T$ is constant, so we could guess that for bosons $T$ must decrease. Since the level spacings shrink generally, the gap between the ground state
and the first excited state also shrinks. This destabilizes the condensate (makes them easier to ‘evaporate’ into non-condensate state). Thus, the total energy increases. Therefore, you must cool the system to keep $E$. That is, $T$ goes down as expected.

Needless to say, under this condition the contribution from the non-condensate is opposite, but its low-lying energy states are much less populated than the ground state below $T_c$, so you must cool the system. More quantitatively, the boson case is easy: Since we may assume $\mu = 0$, we know

$$E = TS - PV = TS - \frac{2}{3}E \Rightarrow E = \frac{3}{5}TS. \quad (22.37)$$

Expanding $V$ reduces the condensate, so $S$ increases. Hence, $T$ must be decreased under constant $E$. In this case we can do better: we can explicitly obtain $E$ from integration

$$E = \int d\varepsilon D_{\varepsilon}(\varepsilon) \frac{\varepsilon}{e^{\beta \varepsilon} - 1} \propto VT^{5/2}, \quad (22.38)$$

so $T$ must be decreased as $V^{-5/2}$. We could use (22.36) which is also correct for free bosons as well. We know $C_V \propto T^\theta$ for some $\theta > 0$ (of course, we know $\theta = 3/2$, but we do not need the exact value), so

$$TC_V - E = \frac{\theta}{1 + \theta}TC_V > 0, \quad (22.39)$$

which is the opposite of fermions.

(2) Since the entropy is constant, we may imagine a situation in which the particles move with the energy levels. However, the level spacings decrease, so excitation would be easier with the initial temperature. To keep $S$ we must maintain the shape of the occupation number distribution. In particular, the condensate population in case of (B) and the cliff shape in case of (F) must be maintained. Hence, $T$ must be decreased. To be quantitative, we need

$$\left. \frac{\partial T}{\partial V} \right|_S = \frac{\partial(T, S)}{\partial(V, S)} = \frac{\partial(T, S)}{\partial(V, P)} \frac{\partial(V, P)}{\partial(V, S)} = -\frac{\partial(V, P)}{\partial(V, S)}$$

$$= -\frac{\partial(V, T)}{\partial(V, S)} \frac{\partial(V, P)}{\partial(V, T)} = -\frac{T}{C_V} \frac{\partial P}{\partial T} \bigg|_V = -\frac{2T}{3V} < 0. \quad (22.40)$$

Here, our favorite $P = 2E/3V$ has been used.

(3) $(\partial P/\partial V)_T < 0$ thermodynamically!.

**Q22-5** Consider a 1D ideal fermion system.

(1) Find the density of one particle states, assuming that the volume (length) is $V$.

(2) Obtain $PV/E$.  

331
(3) Assuming that $E = E_0 + \alpha T^2/2$ for sufficiently small $T$, compute the entropy for sufficiently small $T$.

(4) Find the chemical potential $\mu$ to order $T^2$ using $\alpha$.

**Soln.**

(1) Let us use our usual short cut:

$$\int_{0}^{\varepsilon} d\varepsilon D_t(\varepsilon) = \frac{V}{h} \int_{-\sqrt{2m\varepsilon}}^{\sqrt{2m\varepsilon}} dp = \frac{V}{h} 2\sqrt{2m\varepsilon} = \frac{2\sqrt{2mV}}{h} \varepsilon^{1/2}. \quad (22.42)$$

Therefore,

$$D_t(\varepsilon) = \frac{\sqrt{2mV}}{h} \varepsilon^{-1/2}. \quad (22.43)$$

(2) This implies (or comparing (22.42) and (22.43))

$$\int_{0}^{\varepsilon} d\varepsilon D_t(\varepsilon) = 2\varepsilon D_t(\varepsilon), \quad (22.44)$$

so

$$PV = \mp k_B T \int_{0}^{\varepsilon} d\varepsilon (2\varepsilon D_t(\varepsilon))' \log(1 \mp e^{-\beta(\varepsilon-\mu)}) = 2 \int_{0}^{\varepsilon} d\varepsilon \frac{\varepsilon}{e^{\beta(\varepsilon-\mu)} \mp 1} D_t(\varepsilon) = 2E. \quad (22.45)$$

(3) From the formula for $E$, we get $C_V = \alpha T$. We know

$$\left. \frac{\partial S}{\partial T} \right|_V = \frac{C_V}{T} = \alpha, \quad (22.46)$$

so $S = C_V = \alpha T$, since $S(0) = 0$ (the third law; this is just the same as in 3-space).

(4) We know

$$N\mu = E - TS + PV = 3E - \alpha T^2 = 3E_0 + \frac{1}{2} \alpha T^2 = N\mu_0 + \frac{1}{2} \alpha T^2. \quad (22.47)$$

Notice that $\mu$ increases with $T$. This is the opposite of the case in 3-space.

If you repeat the above calculation in 2-space, you will realize that to order $T^2$, the chemical potential is independent of $T$.

**Warning:** In the above we used a fundamental thermodynamic relation $E = TS - PV + \mu N$. You might think that since we discuss systems with constant $N$, honestly speaking we are using the canonical ensemble, so the basic Gibbs relation is $dE = TdS - PdV$, so $E = TS - PV$ must be correct. Since $PV = 2E/3$, you might write $E = 2TS/5$. The RHS vanishes in the $T \to 0$ limit. However, for the fermion system, we say $E = E_0 + O[T^2]$, and $E_0$ is huge, usually. How come?
If you use the grand canonical ensemble, you must stick to it to be consistent. The energy zero is the zero of one particle state in this formalism (not the ground state energy of the whole system). This causes the discrepancy $E_0$, which is the ground state energy of the whole system relative to the ground state energy of a single particle (for the boson case, there is no discrepancy). That is, with $E = E_0 + O[T^2]$ you must use $E = TS - PV + \mu N$. $PV = 2E/3$ is derived with the aid of the grand canonical formalism, so this $E$ contains $E_0$. The discrepancy occurs only when we discuss the fermion system.

An obvious lesson is that whenever you use $E$ instead of $\delta E$, you must stick to your initial choice of the energy origin.

**Q22-6.** Consider a quantum ideal gas (fermion and boson cases separately, if different). [Hint: Use thermodynamics (as much as possible to save your time) and $PV = 2E/3$ in this problem.]

1. The volume is increased under constant temperature. Does the entropy increase? You must demonstrate your result without any hand-waving argument. Notice that thermodynamics alone cannot answer this question.
2. You wish to decrease the temperature while keeping the pressure. How do you have to change the system volume?

**Soln.**

1. 
   
   $$\frac{\partial S}{\partial V} \bigg|_T = \frac{\partial (S, T)}{\partial (V, T)} = \frac{\partial (S, T)}{\partial (V, -P)} \frac{\partial (V, -P)}{\partial (V, T)} = -\frac{\partial (-P)}{\partial T} \bigg|_V = \frac{\partial P}{\partial T} \bigg|_V = \frac{2}{3V} \frac{\partial E}{\partial T} \bigg|_V > 0.$$

   In this case there is no difference due to statistics.

2. 
   
   $$\frac{\partial V}{\partial T} \bigg|_P = \frac{\partial (V, P)}{\partial (T, P)} = \frac{\partial (V, T)}{\partial (T, P)} \frac{\partial (V, P)}{\partial (V, T)} = -\frac{\partial V}{\partial P} \bigg|_T \frac{\partial P}{\partial T} \bigg|_V > 0,$$

   where we have used the system stability condition. Or more simply
   
   $$\frac{\partial V}{\partial T} \bigg|_P = \frac{\partial (2E/3P)}{\partial T} \bigg|_P = \frac{2}{3P} \frac{\partial E}{\partial T} \bigg|_P = \frac{2C_P}{3V} > 0.$$

   Therefore, the volume must be decreased.

**Q22-7.** Consider a non-interacting boson system below $T_c$ in 3-space. Its internal energy $E$ is proportional to $T^{5/2}$.

1. Show this with the aid of statistical mechanics.
2. Show this using thermodynamics, knowing that $PV = 2E/3$. 

333
[Hint. Simply mimic what we did for the Stefan-Boltzmann law. You need not rederive the formulas derived in the lectures.]

**Soln.**

(1) 

\[
E = \int_0^\infty d\varepsilon \frac{\varepsilon}{e^{\beta\varepsilon} - 1} \propto \int_0^\infty d\varepsilon \varepsilon^{1/2} \frac{\varepsilon}{e^{\beta\varepsilon} - 1} = \beta^{-5/2} \int_0^\infty d(\beta\varepsilon) \left(\frac{\beta\varepsilon}{e^{\beta\varepsilon} - 1}\right)^{1/2},
\]

so indeed \(E \propto T^{5/2}\).

(2) We use \(E = TS - PV\), because \(\mu = 0\). Therefore, we have

\[
S = \frac{5E}{3T}.
\]

Differentiating this wrt \(E\) under constant \(V\), we get

\[
-\frac{2}{3T} = -\frac{5}{3T^2} \frac{\partial T}{\partial E} \bigg|_V E.
\]

or

\[
\frac{dE}{E} = \frac{5}{2} \frac{dT}{T}.
\]

That is, \(E \propto T^{5/2}\).

**Q22-8.** Consider bosons whose dispersion relation is ultrarelativistic, i.e., \(\varepsilon = c|p|\) (as photons).

(1) In 3-space, this system exhibits a Bose-Einstein condensation. What is the number density \(n\) dependence of the critical temperature \(T_c\)? \(\text{Find } \theta \text{ in } T_c \propto n^\theta.\)

(2) Is there any Bose-Einstein condensation in 2-space for these bosons? \([\text{Needless to say, you must justify your opinion.}]\)

**Soln.**

(1) Let us compute the noncondensate population \(N_1\) in 3D:

\[
N_1 = \int d\varepsilon D_t(\varepsilon) \frac{1}{e^{\beta\varepsilon} - 1}. \quad (22.48)
\]

We need the density of (one particle states) \(D_t\). Let us use a shortcut (or you can copy the needed result from somewhere in the lecture notes):

\[
\int_0^\varepsilon D_t(\varepsilon) d\varepsilon = \frac{V}{h^3} \int_{|p|<\varepsilon/c} d^3p = \frac{V}{h^3} \frac{4\pi \varepsilon^3}{3} c^3. \quad (22.49)
\]

334
That is,

\[ D_t(\varepsilon) = \frac{4\pi V}{\hbar^3 c^3 \varepsilon^2}. \]  

(22.50)

Therefore,

\[ N_1 = \int d\varepsilon \frac{1}{e^{\beta \varepsilon} - 1} \propto V T^3 \]  

(22.51)

Therefore, at the critical temperature \( n \propto T_c^3 \), or \( \theta = 1/3 \).

(2) Let us compute the noncondensate population \( N_1 \) in 2D. We need the density of state in 2D:

\[ \int_0^\varepsilon D_t(\varepsilon) d\varepsilon = \frac{V}{\hbar^2} \int_{|p|<\varepsilon/c} d^2p = \frac{V}{\hbar^2} \pi \frac{\varepsilon^2}{c^2}. \]  

(22.52)

That is,

\[ D_t(\varepsilon) = \frac{2\pi V}{\hbar^2 c^2 \varepsilon}. \]  

(22.53)

Now the continuous expression of the number of particles is

\[ N_1 = \int d\varepsilon D_t(\varepsilon) \frac{1}{e^{\beta (\varepsilon - \mu)} - 1} \leq \int d\varepsilon D_t(\varepsilon) \frac{1}{e^{\beta \varepsilon} - 1}. \]  

(22.54)

The question is whether this is finite or not. Since \( D_t \propto \varepsilon \), and for small \( \varepsilon \),

\[ \frac{1}{e^{\beta \varepsilon} - 1} \sim \frac{k_B T}{\varepsilon}. \]  

(22.55)

Therefore, there is no divergence of the integral due the small portion of \( \varepsilon \). The integrand is exponentially small for large \( \varepsilon \), \( N_1 \) is bounded from above by a constant proportional to \( V T^2 \). Therefore, there must be a Bose-Einstein condensation at sufficiently low temperatures.
23 Photons, Phonons and Internal Motions

Summary
* The photon gas statistical thermodynamics is explained.
* We will guess low temperature behaviors of \( E, S, \mu \) for free fermions.
* Analogy to \( \mu = 0 \) grand canonical ensemble may be useful, but do no read it too literally.

Key words
photons, Planck’s radiation formula, ultraviolet catastrophe, Stefan-Boltzmann law, internal degrees of freedom, vibrational and rotational partition functions

What you should be able to do
* You must be able to derive Planck’s formula.
* You must clearly recognize the main features of Planck’s formula.
* You must be able to itemize internal degrees of freedom of a molecule and tell their energy scales (in K).

23.1 Quantization of harmonic degrees of freedom
Photons and phonons are obtained through quantization of the systems that can be described as a collection of harmonic oscillators.\(^{261}\) Possible energy levels for the \( i \)-th mode whose angular frequency is \( \omega_i \)\(^{262}\) are \((n + 1/2)\hbar\omega_i\), where \( n = 0, 1, 2, \cdots \). The canonical partition function of a system with modes \( \{\omega_i\} \) is given by

\[
Z(\beta) = \prod_i \left( \sum_{n_i=0}^{\infty} e^{-\beta(n_i+1/2)\hbar\omega_i} \right),
\]

since no modes interact with 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.

---

\(^{261}\)That is, the system whose Hamiltonian is quadratic in canonical coordinates (quantum mechanically in the corresponding operators).

\(^{262}\)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.
oscillator, which we have already computed. The canonical partition function may be rewritten as:

$$Z(\beta) = \left[ \prod_i \left( e^{-\beta \hbar \omega_i / 2} \right) \right] \left[ \prod_i \left( \sum_{n_i=0}^{\infty} e^{-\beta n_i \hbar \omega_i} \right) \right] \Xi(\beta, 0). \quad (23.2)$$

Here, we have used the formula

$$\Xi(\beta, 0) = \prod_i \left( \sum_{n_i=0}^{\infty} e^{-\beta n_i \hbar \omega_i} \right), \quad (23.3)$$

which may be obtained from the definition of the grand partition function 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 $\sum_i \hbar \omega_i / 2$ is constant and may be ignored by shifting the energy origin.\(^{263}\)

Therefore, the canonical partition function of the system consisting of harmonic modes (or equivalently, consisting of photons or phonons) may be written as $\Xi_{BE}(\beta, 0)$, regarding each mode $\hbar \omega_i$ as a single particle state energy. That is, it is written as the bosonic grand partition function with a zero chemical potential. From this observation, you should immediately recognize that $T$ dependence of various thermodynamic quantities can be computed easily (or dimensional analytic approaches allow us to guess many $T$-dependent behaviors).

23.2 Warning: grand partition function with $\mu = 0$ is only a gimmick

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. Notice that we do not have any more ‘handle’ like $\mu$ to modify the expectation value. 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

\(^{263}\)Warning: 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. For electromagnetic field, the change of the total zero-point energy may be observed as force. This is the Casimir effect.
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 re-
lation is only a mathematical formal relation that can be sometimes useful.\footnote{Intuitively speaking, chemical potential may be defined only for particles you can \textquoteleft pick up.\textquoteright \ More precisely speaking, if no (conserved) charge of some kind (say, electric charge, baryon number) is associated with the particle, its chemical potential is a dubious concept.}

### 23.3 Expectation number of photons

The \( \mu = 0 \) boson analogy tells us that the average number of phonons of a harmonic
mode is given by
\[
\langle n \rangle = \frac{1}{e^{\beta \hbar \omega} - 1}.
\]  
(23.4)

### 23.4 Internal energy of photon systems

The phonon contribution to the internal energy of a system may be computed just
as we did for the Debye model. We need the density of states (i.e, phonon spectrum,
i.e., the distribution of the frequencies of the modes) \( D_{\text{ph}}(\omega) \). The internal energy
of all the phonons is given by
\[
E = \sum_{\text{modes}} \langle n(\omega) \rangle \hbar \omega = \int d\omega D_{\text{ph}}(\omega) \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}.
\]  
(23.5)

This is the internal energy without the contribution of zero-point energy.

A standard way to obtain the density of states \( D_{\text{ph}}(\omega) \) is to study the wave
equation governing the electromagnetic waves, but here we use our usual shortcut.
The dispersion relation for photons is \( \varepsilon = c |p| = \hbar \omega \), so
\[
\int_0^\omega D_{\text{ph}}(\omega') d\omega' = \frac{V}{\hbar^3} \int_{|p| \leq \hbar \omega / c} d^3 p.
\]  
(23.6)

Here, we do not include the factor 2 due to polarization states. Differentiating the
above equality, we obtain
\[
D_{\text{ph}}(\omega) = \frac{4\pi V}{\hbar^3} \left( \frac{\hbar \omega}{c} \right)^2 \frac{\hbar}{c} = \frac{V \omega^2}{2 \pi^2 c^3}.
\]  
(23.7)
Photons have two polarization directions, so the actual density of the modes is this formula $\times 2$.

### 23.5 Planck’s distribution, or radiation formula

The internal energy $dE_\omega$ and the number $dN_\omega$ of photons in $[\omega, \omega + d\omega)$ in a box of volume $V$ are given by

$$dE_\omega = 2D_{ph}(\omega)\frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}d\omega,$$

$$dN_\omega = 2D_{ph}(\omega)\frac{1}{e^{\beta \hbar \omega} - 1}d\omega.$$

The factor 2 comes from the polarization states (i.e., $D_{ph}$ here is given by (23.7)).

Therefore, the energy density $u(T, \omega)$ at temperature $T$ due to the photons with the angular frequencies around $\omega$ reads

$$u(T, \omega) = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}.$$  

This is Planck’s radiation formula.

![Figure 23.1: Classical electrodynamics gives the Rayleigh-Jeans formula (23.12) (green); this is the result of equipartition of energy and due to many UV modes, the density is not integrable (the total energy diverges). Wien reached (23.13) empirically (red). Planck arrived at his formula (black) originally by interpolation of these two results. Notice that the peak position is proportional to the temperature. It is important to know some qualitative features of this law (Fig. 23.1): (i) Planck’s law can explain why the spectrum blue-shifts as temperature increases; this was not possible within the classical theory. (ii) The total energy density $u(T) = E/V$ of a radiation field at temperature $T$ is finite. $u(T)$ is obtained by integration:

$$u(T) = \int_0^{\infty} d\omega u(T, \omega).$$  

---

---

---
With Planck’s law (23.10) this is always finite (we will study this later).

(iii) In the classical limit $\hbar \to 0$, we get

$$u(T, \omega) = \frac{k_B T \omega^2}{\pi^2 c^3} \left( = 2D_{ph}(\omega)k_B T \right),$$

(23.12)

which is the formula obtained by classical physics (i.e., the equipartition of energy). Upon integration, the classical limit gives an infinite $u(T)$. This divergence is obviously due to the contribution from the high frequency modes. Thus this difficulty is called the ultraviolet catastrophe, which destroyed classical physics.

(iv) In the high frequency limit $\hbar \omega \gg k_B T$ Planck’s law (23.10) goes to

$$u(T, \omega) \approx \frac{k_B T}{\pi^2 c^3} \omega^2 e^{-\beta \hbar \omega},$$

(23.13)

which was empirically proposed by Wien.

### 23.6 Statistical thermodynamics of black-body radiation

Let us finish the statistical mechanics of black-body radiation.

$$u(T) = \int_{0}^{\infty} \frac{\omega^2}{\pi^2 c^3} \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} d\omega = \beta^{-4} \int_{0}^{\infty} \frac{(\beta \omega)^2}{\pi^2 c^3} \frac{h \beta \omega}{e^{\beta \hbar \omega} - 1} d(\beta \omega).$$

(23.14)

This immediately implies (as seen above)

$$u(T) \propto T^4,$$

(23.15)

which is called the Stefan-Boltzmann law.\(^{266}\)

Since we know the $T^3$-law of the phonon low temperature specific heat (the Debye theory), this should be expected. This is understandable by counting the number of degrees of freedom (Fig. 22.1) explained before. Although we did not calculate the proportionality constant, if you follow the above calculation you can get it. This proportionality was obtained purely thermodynamically by Boltzmann before the advent of quantum mechanics. The proportionality constant contains $\hbar$, so it was impossible to theoretically obtain the proportionality constant before Planck (Stefan experimentally obtained it).

----

\(^{266}\)The proportionality constant can be computed as $k_B^4 \pi^2 / 15 \hbar^3 c^3$. 

340
23.7 Black-body equation of state
Photons may be treated as ideal bosons with \( \mu = 0 \), so the equation of state is immediately obtained as

\[
\frac{PV}{k_B T} = \log \Xi = - \int d\varepsilon D(\varepsilon) \log(1 - e^{-\beta \varepsilon}). \tag{23.16}
\]

For 3D superrelativistic particles, \( D(\varepsilon) \propto \varepsilon^2 \), so

\[
\int_0^\varepsilon d\varepsilon D(\varepsilon) = \frac{1}{3} \varepsilon D(\varepsilon). \tag{23.17}
\]

This gives us (review what we did to derive \( PV = 2E/3 \) for ordinary particles in 21.15)

\[
PV = \frac{1}{3} E. \tag{23.18}
\]

23.8 Thermodynamic derivation of black-body equation of state
Just as \( PV = 2E/3 \) is a result of pure mechanics, (23.18) is a result of pure electrodynamics, so this was known before quantum mechanics. Boltzmann started with (23.18) to obtain the Stefan-Boltzmann law as follows.

Since we know generally

\[
E = TS - PV = TS - \frac{1}{3} E, \tag{23.19}
\]

\[
ST = \frac{4}{3} E \quad \text{or} \quad S = \frac{4}{3} \frac{E}{T}. \tag{23.20}
\]

Differentiating \( S \) wrt \( E \) under constant \( V \), noting \( (\partial S/\partial E)_V = 1/T \), we obtain

\[
\frac{1}{T} = - \frac{4}{3T^2} \frac{\partial T}{\partial E}_V \bigg|_V E + \frac{4}{3T} \tag{23.21}
\]

or

\[
\frac{1}{3T} = \frac{4}{3T^2} \frac{\partial T}{\partial E}_V \bigg|_V E, \tag{23.22}
\]

that is, under constant \( V \)

\[
\frac{dE}{E} = 4 \frac{dT}{T}. \tag{23.23}
\]

If \( \mu = 0 \), then \( A = -PV = -k_B T \log Z \).

341
This implies the Stefan-Boltzmann law \( E \propto T^4 \). The proportionality coefficient contains \( \hbar \), so Boltzmann could not get it; Stefan experimentally determined the value.

### 23.9 Blackbody - low temperature phonon system analogy

For a phonon system of a lattice, we have a high-frequency cutoff in the energy spectrum, but its effect is almost negligible in the low temperature limit. Except for the number of modes, you must clearly recognize a direct relation between the photons in the vacuum and phonons in the crystal. Debye’s \( T^3 \) law is ‘almost the same’ as the Stefan-Boltzmann law.

### 23.10 Internal degrees of freedom of classical ideal gas

If noninteracting particles are sufficiently dilute (\( \mu \ll 0 \)), we know classical ideal gas approximation is OK. However, the internal degrees of freedom may not be handled classically, because energy gaps may be huge. We have already glimpsed this when we discussed the gas specific heat.

Let us itemize internal degrees of freedom of a molecule:

i) Each atom has a nucleus, and its ground state could have nonzero nuclear spin. This interacts with electronic angular momentum to produce the *ultrafine structure*. The splitting due to this effect is very small, so for the temperature range relevant to the gas phase we may assume all the levels are energetically equal. Thus, (usually) we can simply assume that the partition function is multiplied by a constant \( g \) = degeneracy of the nuclear ground state.\(^{268}\)

ii) Electronic degrees of freedom has a large excitation energy (of order of ionization potential \( \sim \) a few eV, so unless the ground state of the orbital electrons is degenerate), we may ignore it.\(^{269}\)

iii) If a molecule contains more than one atom, it can exhibit rotational motion. The quantum of rotational energy (\( \Theta_R \) below) is usually of order 10 K.\(^{270}\)

iv) Also such a molecule can vibrate. The vibrational quantum (\( \Theta_V \) below) is of

\(^{268}\)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 as is stated in the text.

\(^{269}\)If the ground state is degenerate, then it could have a fine structure with an energy splitting of order a few hundred K. For ground state oxygen (^3P_2) the splitting energy is about 200 K, so we cannot simply assume that all the states are equally probable nor that only the ground state is relevant.

\(^{270}\)However, for H\(_2\) it is 85.4 K. For other molecules, the rotational quantum is rather small: N\(_2\): 2.9 K; HCl: 15.1 K.
order 1000 K.\textsuperscript{271}

\subsection*{23.11 Rotation and vibration}
Notice that there is a wide temperature range, including the room temperature, where we can ignore vibrational excitations and can treat rotation classically (Fig. 23.2). Thus, equipartition of energy applied to translational and rotational degrees of freedom can explain the specific heat of many gases.

![Figure 23.2: The constant volume specific heat.](image)

The Hamiltonian for the internal degrees of freedom for a diatomic molecule reads

\[ H = \frac{1}{2} J^2 + \hbar \omega \left( \hat{n} + \frac{1}{2} \right), \]  

(23.24)

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 \( z_i = z_r z_v \):

\[ z_r = \sum_{J=0}^{\infty} (2J + 1) e^{-(\Theta_R/T)J(J+1)}, \]  

(23.25)

with \( \Theta_R = \hbar^2 / 2k_B I \) and

\[ z_v = \sum_{n=0}^{\infty} e^{-(\Theta_V/T)(n+1/2)}, \]  

(23.26)

with \( \Theta_V = \hbar \omega / k_B \).

\textsuperscript{271}N\textsubscript{2}: 3340 K; O\textsubscript{2}: 2260 K; H\textsubscript{2}: 6100 K.
23.12 Low and high temperature limit of rotational contribution

If the temperature is sufficiently low, then

\[ z_r \simeq 1 + 3 e^{-2\Theta_R/T}. \] (23.27)

The contribution of rotation to specific heat is

\[ C_{\text{rot}} \simeq 3Nk_B \left( \frac{\Theta_R}{T} \right)^2 e^{-2\Theta_R/T}. \] (23.28)

For \( T \gg \Theta_R \), we may approximate the summation by integration (Large \( J \)s contribute, so we may approximate \( J \simeq J + 1 \)):

\[ z_r \simeq 2 \int_0^{\infty} dJ J e^{-J^2(\Theta_R/T)} = \frac{T}{\Theta_R}. \] (23.29)

This gives the rotational specific heat, but it is more easily obtained by the equipartition of energy, because the rotational energy is with a quadratic form. Thus, \( C_{\text{rot}} = k_B \) in the high temperature limit.

23.13 Low and high temperature limit of vibrational contribution

The vibrational partition function can be summed as

\[ z_v = 1/2 \sinh(\beta \hbar \omega / 2). \] (23.30)

For small \( T \)

\[ z_v \simeq (1 + e^{-\beta \omega})e^{\beta \omega / 2} \] (23.31)

is enough. Consequently,

\[ C_{\text{vib}} \sim k_B N \left( \frac{\Theta_R}{T} \right)^2 e^{-\Theta_R/T}. \] (23.32)

Since \( \Theta_R \ll \Theta_V \), as already noted, there is a wide range of temperature where only rotation contributes to the specific heat.
24 Phases and phase transitions

Summary
* Statistical thermodynamics is briefly reviewed with illustrations relevant to the current topic.
* Qualitative change of phases is the phase transition, which corresponds to some singularities of thermodynamic potentials.
* Phase coexistence conditions (under given $T$ and $P$) are a set of equalities among chemical potentials. Gibbs’ phase rule follows from the condition.
* Thermodynamic limit is absolutely needed to rationalize phase transitions statistically.

Key words
phase, phase transition, phase diagram, coexistence curve, triple point, kelvin scale, Gibbs’ phase rule, first order phase transition, second order phase transition, Ising model, thermodynamic limit.

What you should be able to do
* You must be able to draw the phase diagram of a ordinary one-component fluid on the PT plane.
* You must be able to sketch $G(T, P)$ for an ordinary fluid.

So far we have not discussed systems with interactions. If there are interactions, there are various phases as is exemplified by ice, liquid and vapor of water. First, let us discuss how to describe thermodynamically what we experience. Then, let us discuss whether statistical mechanics can discuss phase transitions.

24.1 What is a phase?
Intuitively, under different conditions (say, at various $(T, P)$) a system can exhibit qualitatively different properties. When this happens, we say the system (or the material) is in different phases.\textsuperscript{272} To understand a substance is to understand its various phases and their characteristic features. Therefore, we wish to map out what happens at various points in the thermodynamic space or at least in terms of

\textsuperscript{272} We may say an equilibrium state is in a single phase, if it is macroscopically homogeneous.
thermodynamic parameters (e.g., $T$, $P$, etc.), i.e., we wish to construct the phase diagram (see, e.g., Fig. 24.1). To understand the world we must understand where the state boundaries are and what the features of the territories are. To understand the boundaries corresponds to the understanding of phase transitions, and to understand the features of the territories corresponds to characterizing individual phases.

In the above, ‘qualitative differences’ do not imply quantitative difference such as soft-hard, hot-cold, hue changes, etc., but existence-non existence of some properties such as symmetry, long-range correlation, etc. For example, solid, liquid, and gas phases may be characterized by the following table:

<table>
<thead>
<tr>
<th></th>
<th>long-range order</th>
<th>coherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>solid</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>liquid</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>gas</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Here, ‘long-range’ correlation implies that if you know a position of a particle, you can tell the position of another particle far away from the first one. Crystalline spatial regularity implies long-range spatial ordering of the particles. For fluid phases we cannot have this property. This property can either exist or not exist. This is the qualitative difference between solid and fluid phases. To distinguish fluid phases is not easy. One possibility is stated in the above table. We know gases can be compressed easily but liquids cannot; they are as incompressible as solids. This must be due to the interactions (‘touching’) among molecular hard cores. ‘coherence’ implies that each particle has at least four repulsive interactions with its surrounding particles simultaneously.

Since some qualitative properties appear or disappear upon crossing a phase boundary, something ‘singular’ can happen thermodynamically (e.g., loss of differentiability or continuity of some thermodynamic quantities), and this change is called a phase transition.

![Figure 24.1: A representative phase diagram of an ordinary fluid. S: solid; L: liquid; G: gas; t: triple point; cp: critical point. The curves denote the phase boundaries where phase transitions occur.](image-url)
24.2 Phase may not sometimes be well-defined globally
However, a precise definition of ‘phase’ is actually rather difficult. Near the phase boundaries we may clearly distinguish the phases, but the ‘territory where a phase occupies’ may not be well-defined as in the case of gases and liquids. Therefore, here, 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) are distinct phases (near the phase transition).

Since we have come a long way from Maxwell, Clausius, Boltzmann and other founding fathers of our subject, let us review salient points of statistical thermodynamics that we need to understand phase transitions with relevant illustrations.

24.3 Phase diagram in thermodynamic space
For a given system its any equilibrium state is described (uniquely) as a point in its thermodynamic space spanned by its thermodynamic coordinates (= its internal energy $E$ and work coordinates $X$, say, $V$). Thermodynamic coordinates are crucial when we build thermodynamics, since they do not require thermodynamics to describe (purely mechanical description is possible). The unique relation of a point in the thermodynamic space to an equilibrium state of the system allows us to describe phase coexistence unambiguously.

To illustrate this point look at the ordinary solid-liquid-gas phase diagram (Fig. 24.2).

24.4 Thermodynamic potentials and partition functions
You must be able to write the Gibbs relation
\[ dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{B}{T}dM - \frac{\mu}{T}dN + \cdots, \]
(24.1)
and the corresponding Gibbs-Duhem relation
\[ Ed\frac{1}{T} + Vd\frac{P}{T} - Md\frac{B}{T} - Nd\frac{\mu}{T} + \cdots = 0. \]
(24.2)
The fundamental principle of statistical thermodynamics is the following translation of entropy into phase volume compatible with a point in the thermodynamic space by Boltzmann:

\[ S = k_B \log w(E, X). \] (24.3)

However, we often wish to describe the phase transition under constant \( T, P, \) etc., so we must Legendre transform \( S \) to a generalized Gibbs free energy \( \tilde{G} \).

\[ S \rightarrow S - \frac{1}{T}E - \frac{P}{T}V + \frac{B}{T}M = -\frac{\tilde{G}}{T} \] (24.4)

The generalized canonical formalism reads

\[ \tilde{G} = -k_B T \log \tilde{Z} \text{ with } \tilde{Z} = \sum_{E,V,M} \omega(E, V, M)e^{-\beta(E+PV-BM)}. \] (24.5)

### 24.5 Stability, fluctuation and response

The equilibrium condition under \( T, P, \) \( \cdots \) constant is the minimization of \( \tilde{G} \). However, the equilibrium stability condition can always be written as \( \delta^2 E \geq 0 \) (or \( \delta^2 S \leq 0 \)). The positivity of the diagonal terms of this Hessian is le Chatelier’s principle

\[ \left. \frac{\partial X}{\partial x} \right|_y > 0, \] (24.6)

\footnote{Generally, the Legendre transformation of entropy is called \textit{Massieu functions}. \(-\tilde{G}/T\) is a typical example.}
where $y$ denotes other intensive quantities. Here, we have assumed that $E$ is twice differentiable, but this does not always hold. For example, if a phase transition occurs, this is highly questionable. Recall that the really important point is that $E$ is a convex function. It is guaranteed that $E$ is continuously once differentiable (i.e., a $C_1$ function), but the second derivatives may not exist. The true implication of (24.6) is that if $x$ increases, then so does $X$.

Suppose a phase transition between phase I and II occurs. If the transition to II occurs from I by increasing temperature (i.e., II is a higher temperature phase than I), then $S_I < S_{II}$, because increasing $T$ implies increasing its conjugate variable: entropy. Or, if II is a higher pressure phase than I, then $-V_{II} > -V_I$, i.e., $V_I > V_{II}$ (the conjugate variable of $P$ is not $V$ but $-V$).

We discussed the fluctuation-response relation

$$\chi = \frac{\partial X}{\partial x} \bigg|_y = \beta \langle \delta X^2 \rangle. \quad (24.7)$$

If the ordered phase is stable, to destroy its order very large fluctuation is needed. This actually happens near the critical point, and for example magnetic susceptibility or compressibility diverges.

24.6 Can phases coexist?
If there is a phase transition from one phase to another, there may or may not be a coexistence of these phases.\textsuperscript{274} Thus, ‘condition’ in the following means, at best, a necessary condition.

24.7 Coexistence condition for two phases
Suppose a system is described by its thermodynamic coordinates $(E, X)$ and two phases I and II coexist in equilibrium when $E$ and $X$ exchange is allowed between the phases (in an isolated box). When two phases coexist, the substance may be exchanged freely across the phase boundary between the two phase. Therefore, we must also take the exchange of the material into account. We know $S = S_I + S_{II}$ must be maximized, so the equilibrium condition is the identity of $T$, $x/T$ and $\mu/T$ as we discussed long ago. Let us review it. The Gibbs relation implies

$$\delta S = \frac{1}{T} \delta E - \frac{x}{T} \delta X - \frac{\mu}{T} \delta N. \quad (24.8)$$

\textsuperscript{274}It is impossible to know whether the given two phases coexist or not purely thermodynamically, although, usually, we may say the phases satisfying the thermodynamic coexistence conditions do coexist.
Here, $\delta$ implies variation or virtual change, BUT in reality fluctuations actually realize these needed changes spontaneously. $\delta E_I + \delta E_{II} = 0$, $\delta X_I + \delta X_{II} = 0$ and $\delta N_I + \delta N_{II} = 0$ imply the following equilibrium condition:

$$
\delta S = \frac{1}{T_I} \delta E_I - \frac{x_I}{T_I} \delta X_I - \frac{\mu_I}{T_I} \delta N_I + \frac{1}{T_{II}} \delta E_{II} - \frac{x_{II}}{T_{II}} \delta X_{II} - \frac{\mu_{II}}{T_{II}} \delta N_{II} = 0. \tag{24.9}
$$

Therefore, $T_I = T_{II}$ and $x_I = x_{II}$ are required in general. Usually, $X = V$ and $N$, so $T, P$ and $\mu$ must be identical between two phases:

$$
T_I = T_{II}, \quad P_I = P_{II}, \quad \mu_I = \mu_{II}. \tag{24.10}
$$

The last equality in (24.11) is

$$
\mu_I(T, P) = \mu_{II}(T, P). \tag{24.12}
$$

This functional relation determines a curve called the coexistence curve in the $T$-$P$ diagram (see Fig. 24.1).\textsuperscript{275}

Along this line the Gibbs free energy $G$ of the whole system may be written as

$$
G = N_I \mu_I + N_{II} \mu_{II}. \tag{24.13}
$$

Thus, without changing the value of $G$, any mass ratio of the two phases is admissible, if they can coexist. This implies that in the phase diagram in the thermodynamic space the phase coexistence relation is described as a boundary ‘mine field’ (see Fig. 24.2 Right) instead of a line.

**24.8 Phase coexistence condition: pure substance case**

How many phases can coexist at a given $T$ and $P$? Suppose we have $X$ coexisting phases. The following conditions must be satisfied:

$$
\mu_I(T, P) = \mu_{II}(T, P) = \cdots = \mu_X(T, P). \tag{24.14}
$$

We believe that for the generic case, $\mu$’s are sufficiently functionally independent. To be able to solve for $T$ and $P$, we can allow at most two independent relations. That

\textsuperscript{275}(24.12) may be obtained by minimizing the Gibbs free energy under the constant $T$ and $P$, because in the thermodynamic limit (i.e., if the system is large enough) what we can obtain from an isolated system and the same system under $T, P$ constant condition with the consistent $T$ and $P$ are indistinguishable (ensemble equivalence).
is, at most three phases can coexist at a given $T$ and $P$ for a pure substance.

For a pure substance, if three phases coexist, $T$ and $P$ are uniquely fixed. This point on the $T$-$P$ diagram is called the *triple point*. The Kelvin scale of temperature is defined so that triple point of water is at $T = 273.16 \text{K}$ (since 1954). $t = T - 273.15$ is the temperature in Celsius. Again, this is the definition of °C.

For a pure substance, as we have seen, the chemical potentials of coexisting phases must be identical. Before and after the phase transition from phase I to II or vice versa, there is no change of the Gibbs free energy

$$\Delta G_{\text{CC}} = 0,$$

where CC means "along the coexistence curve" and $\Delta$ implies the difference across the coexistence curve (say, phase I — phase II). From (24.15) we already obtained the *Clapeyron-Clausius relation*:

$$\left. \frac{\partial P}{\partial T} \right|_{\text{CC}} = \frac{\Delta_{I \rightarrow II} H}{T \Delta_{I \rightarrow II} V},$$

where $\Delta_{I \rightarrow II} X$ denotes $X^{\text{II}} - X^{\text{I}}$.

### 24.9 Gibbs’ phase rule

Consider a more general case of a system consisting of $c$ chemically independent components (i.e., the number of components we can change independently). For example, $\text{H}_3\text{O}^+$ in pure water should not be counted, if we count $\text{H}_2\text{O}$ among the independent chemical components.

Suppose there are $\phi$ coexisting phases. The equilibrium conditions are:

1. $T$ and $P$ must be common to all the phases,
2. The chemical potentials of the $c$ chemical species must be common to all the phases.

To specify the composition of a phase we need $c - 1$ variables, because we need only the concentration ratios. Thus, the chemical potential for a chemical species depends on $T$, $P$ and $c - 1$ mole fractions ($x^1, x^2, \cdots, x^{c-1}$), which are not necessarily common to all the phases (we must add a suffix to denote the phases). That is, $\mu$’s are $c + 1$ variable functions, and we have $2 + \phi(c - 1)$ unknown variables. We have $\phi - 1$ equalities among the chemical potentials in different phases for each chemical species, so the number of equalities we have is $(\phi - 1) \times c$. Look at the following simultaneous equations:

$$\mu_1^{\text{I}}(T, P, x_1^1, x_1^2, \cdots x_1^{c-1}) = \mu_1^{\text{II}}(T, P, x_{\text{II}}^1, x_{\text{II}}^2, \cdots x_{\text{II}}^{c-1}) = \cdots = \mu_\phi^{\text{I}}(T, P, x_\phi^1, x_\phi^2, \cdots x_\phi^{c-1}),$$

351
\[ \mu_1(T,P,x_1^1,x_1^2,\ldots,x_1^{c-1}) = \mu_2(T,P,x_1^1,x_2^2,\ldots,x_2^{c-1}) = \ldots = \mu_\phi(T,P,x_1^1,x_\phi^2,\ldots,x_\phi^{c-1}), \]
\[ \mu_1^I(T,P,x_1^1,x_2^2,\ldots,x_1^{c-1}) = \mu_2^I(T,P,x_1^1,x_2^2,\ldots,x_1^{c-1}) = \ldots = \mu_\phi^I(T,P,x_1^1,x_\phi^2,\ldots,x_\phi^{c-1}), \]
\[ \mu_1^\psi(T,P,x_1^1,x_2^2,\ldots,x_1^{c-1}) = \mu_2^\psi(T,P,x_1^1,x_2^2,\ldots,x_1^{c-1}) = \ldots = \mu_\phi^\psi(T,P,x_1^1,x_\phi^2,\ldots,x_\phi^{c-1}), \]

Consequently, for the generic case we can choose \( f = 2 + \phi(c-1) - c(\phi-1) = c + 2 - \phi \)
variables freely. This number \( f \) is called the number of thermodynamic degrees of freedom. We have arrived at the Gibbs phase rule:\[ \text{24.17} \]
\[ f = c + 2 - \phi. \]  

### 24.10 How \( G \) behaves at the phase boundaries?

What happens to the Gibbs free energy at the phase transition point under constant \( T \) and \( P \)? You must be able to sketch it in the ordinary fluid case. Note the usual Gibbs relation:
\[ dG = -SdT + VdP. \]  

(24.19)

Under constant \( P \), \( G \) may be sketched as follows:

![Typical behavior of Gibbs free energy for a pure substance. The free energy loses differentiability at first order phase transition points.](image)

When \( P \) is the critical pressure, then the liquid-gas transition ‘break’ disappears: \( G \) becomes differentiable. However, the specific heat has a singular behavior at the critical temperature. That is, the LG transition becomes second order. We will discuss this later in more detail.

\[ \text{276As astute readers have probably sensed already, the derivation is not water tight. We have assumed that there is no special functional relations among chemical potentials. Rigorously speaking, we cannot guarantee this and so we cannot derive the phase rule from the fundamental laws of thermodynamics (nor equilibrium statistical mechanics, either).} \]

352
Try to sketch $G$ under constant $T$ as a function of $P$.

### 24.11 Classification of phase transitions

Usually, phase transitions are classified into first order phase transitions and the rest called continuous phase transitions or second-order phase transitions. In the first order phase transition at least one thermodynamic density (= extensive quantity per volume) changes discontinuously, but in the second order phase transitions there is no discontinuity in thermodynamic densities. The liquid-gas transition at the critical pressure is a second-order phase transition as noted just above.

Phase transitions in many interesting cases occur between more ordered and less ordered phases; it is between the low entropy state and the high energy (enthalpy) state. For example melting is the transition from low entropy solid to high energy liquid. Protein folding is the transition from higher energy random coil state to low entropy folded state.  

A first order phase transition occurs if the ordered phase loses its stability ‘catastrophically.’ In other words, the first order phase transition occurs when a slight loss of order favors further loss of order. Thus, there is no equilibrium state with reduced stability. In contrast, in the case of second order phase transitions reduction of order does not appreciably destabilize the order further. Thus, a phase with reduced stability of order can exist as an equilibrium state. You can intuitively understand a stable phase with reduced stability as an oscillator with a very weak spring. Fluctuation becomes very large near the second order phase transition; despite large thermal fluctuations the phase persists. The ordered phase persists until fluctuation becomes indefinitely large. Since the ordering of some sort is always the reason for phase transitions, the second order phase transition is theoretically the most interesting. Therefore, we discuss the second order phase transition first.

### 24.12 Typical example of second order phase transition

A typical second order phase transition is the one from the paramagnetic to the ferromagnetic phase we can observe in magnets.

A magnet can be understood as a lattice of spins interacting with each other locally in space. The interaction between two spins has a tendency to align them parallelly. At higher temperatures, due to vigorous thermal motions, this interaction

---

277However, do not have a prejudice that natural states of proteins are equilibrium states. Many large proteins are likely to be in metastable states when they function biologically normally. Think how you can prove experimentally that a particular protein is in equilibrium.
cannot quite make order among spins, but at lower temperatures the entropic effect becomes less significant, so spins order globally. There is a special temperature $T_c$ below which this ordering occurs. We say an order-disorder transition occurs at this temperature.

The Ising model is the simplest model of this transition. At each lattice point is a (classical) spin $\sigma$ which takes only $+1$ (up) or $-1$ (down). A nearest neighbor spin pair has the following interaction energy:

$$-J\sigma_i\sigma_j,$$

where $J$ is called the coupling constant, which is positive in our example (ferromagnetic case; if spins are parallel, interaction energy is lowered). We assume all the spin-spin interaction energies are superposable, so the total energy of the system for a lattice is given by

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J\sigma_i\sigma_j - \sum_i h\sigma_i,$$

where $\langle \rangle$ implies the nearest neighbor pairs, and $h$ is the external magnetic field.

The (generalized canonical) partition function for this system reads

$$Z = \sum_{\{\sigma_i = \pm1\}} e^{-\beta\mathcal{H}}.$$  

Here, the sum is over all spin configurations.$^{278}$

### 24.13 Fundamental questions about phase transitions

So far phenomenologically, we accepted the existence of phase transitions, and discussed how to describe/handle them. From the statistical mechanics point of view, the most important question is why such qualitative changes can occur at all. Actually, a more fundamental question is: can statistical mechanics ever describe phase transitions? Up to early 1930s such doubts existed. Now, in the 21st century, we are sure that statistical mechanics correctly describes various phase transitions. HOWEVER, do not forget that we cannot yet explain statistical mechanically why

---

$^{278}$What is $F = -k_B T \log Z$ in this case? It is NOT the Helmholtz free energy in the original sense. That is why $Z$ is (precisely speaking) called the generalized canonical partition function. Notice that $dF$ is NOT given by $dF = -SdT + hdM$ (assuming the volume to be constant), but by $dF = -SdT - Mdh$.

Why does this happen? This is due to the term $-\sum h\sigma_i$ in the Hamiltonian (24.21). This is not a part of the proper energy of the system, but the potential energy of the spins stored between the magnet and the device creating the magnetic field.
ordinary molecules can make crystals below some finite temperature. Does the partition function contain a crystal? We believe so, but no one has demonstrated this.

24.14 Necessity of thermodynamic limit
If the system size is finite, the sum in (24.22) is a finite sum of positive terms. Each term in this sum is analytic\(^{279}\) in \(T\) and \(h\), so the sum itself becomes analytic in \(T\) and \(h\) (i.e., very smooth). Furthermore, \(Z\) cannot be zero, because each term in the sum is strictly positive. Therefore, its logarithm is analytic in \(T\) and \(h\); the free energy of the finite lattice system cannot exhibit any singularity. That is, there is no thermodynamic singularity, and consequently, there is no phase transition for this system.\(^{280}\)

Even in the actual system we study experimentally, there are only a finite number of atoms, but this number is huge. Thus, the question of phase transitions from the statistical mechanics point of view is: is there any singularity in \(A = -k_B T \log Z\) in the large system limit? The large system limit, with proper caution not to increase its surface area more than the order of \(V^{2/3}\), where \(V\) is the system volume, is called the thermodynamic limit. Strictly speaking, phase transitions can occur only in this limit.\(^{281}\)

You may not wish to go into mathematics, but at least clearly recognize that qualitative changes require loss of analyticity.

\(^{279}\)A more accurate mathematical term is ‘holomorphic.’

\(^{280}\)Strictly speaking, there is no phase transition for any finite system, unless each spin has infinitely many states.

\(^{281}\)Does such a limit exist? This is also a fundamental question never considered till 1950s. This is a far easier question than the existence/non-existence question of phase transitions.
Q24-1. Consider an Ising magnetic system in 3-space. There is a second order order-disorder phase transition at $T = T_c$, if the magnetic field $h = 0$ (i.e., $h_c = 0$).

1. What is the thermodynamic space for this system (or what are the thermodynamic coordinates for the system)?
2. Sketch the phase diagram of this magnet in its thermodynamic space.

**Soln.**

1. $E$ and $M$ must be the thermodynamic coordinates. $h$ is intensive, and corresponds to $P$ in fluids.
2. We must draw a phase diagram on $(E, M)$. The usual diagram is on $(h, T)$.

![Phase Diagram](image)

Figure 24.4: Below $T_c$ $M \neq 0$ is possible with $h = 0$. If you apply nonzero $h$, one of the up or down phases remain as an equilibrium state. Thus, the ‘lined’ region is with $h = 0$; it is realizable if $d > 2$, but not in $d = 2$ (i.e., no actual equilibrium system can be in this region). Outside of this ‘coexistence’ region, for higher $T$ of the temperature axis, magnetic field must be applied. Needless to say too large $|M|$ is not realizable, so the thermodynamic space is bounded vertically.

The rough sketch is shown here. However, accurate sketch is very hard.

Q24-2. The melting temperature of tetrachlorocarbon (CCl$_4$) depends on the pressure as follows:

$$T_m = 250.56 + 4.005 \times 10^{-2}P - 2.15 \times 10^{-6}P^2,$$

(24.23)

where $T$ is measured in K and $P$ in atm. At $P = 1000$ atm, the melting causes the volume increase of $\Delta V = 3.06 \text{ m}\ell$ per mole. Find the latent heat of melting per mole of tetrachlorocarbon at 1000 atm. Notice that 1 atm = 101,325 P.

**Soln.**

We use the Clapeyron-Clausius relation

$$\frac{\partial T_m}{\partial P} \bigg|_{CC} = \frac{T_m \Delta V}{\Delta H}$$
or

\[ \Delta H = T_m \Delta V \left( \left. \frac{\partial T_m}{\partial P} \right|_{CC} \right)^{-1} \]

\[ \left. \frac{\partial T_m}{\partial P} \right|_{CC} = 4.005 \times 10^{-2} - 4.30 \times 10^{-6} P = 0.0358 \]

in K/atm. Therefore, \((T_m = 288.5\) K\)

\[ \Delta H = 288.5 \times (3.06 \times 10^{-6}/0.0358) \text{ m}^3 \cdot \text{atm} = 0.02466 \text{ m}^3 \cdot \text{atm} = 2498 \text{ J}, \]

or 597 cal/mol.
25 Spatial dimensionality and interaction range are crucial

Summary
* Statistical mechanics seems to be able to explain various phases and phase transitions rationally in the thermodynamic limit.
* Spatial dimensionality is crucial to the phase ordering and existence of the order-disorder phase transition. Peierls’ argument tells us the importance of spatial dimensionality.
* If the interaction is long-ranged, phase transitions can happen even in 1-space. (Augmented) van der Waals gas in 1D is a typical example.
* The second order phase transition for magnets, fluids and binary fluid mixtures may be understood in a unified fashion.

Key words
Peierls’ argument, Kac potential, van der Waals gas, Maxwell’s rule, Tonks’ gas

What you should be able to do
* Intuitively understand why spatial dimensionality matters.
* You must be able to explain Peierls’ argument.
* Derivation of Tonk’s equation of state should be a good exercise.

Whether statistical mechanics can understand phase transitions or not in the thermodynamic limit is a fundamental question. Peierls definitely settled the issue by demonstrating that the 2D Ising model has an ordered phase. His demonstration makes it clear that spatial dimensionality is crucial.

25.1 Magnet-lattice gas correspondence
The Ising model due to Lenz was introduced in the preceding lecture, whose Hamiltonian reads

\[ H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i. \]  

(25.1)

We can interpret this as a lattice model of a gas (lattice gas) with the following correspondence ‘down’ (resp., ‘up’) \( \rightarrow \) lattice point ‘occupied’ (resp., ‘empty’). \( h \) may be interpreted as the chemical potential (large positive \( h \) implies more down spins =
25.2 Order parameter
To characterize the order in the system we define an order parameter which is nonzero only in the ordered phase: magnetization per particle \( m = \langle \sigma \rangle = M/N = (1/N) \sum \sigma_i \) is a good example. Thus, the fundamental question about the Ising model is whether \( M/N \) converges to zero or not in the thermodynamic limit.\(^{282}\)

You can watch 2-Ising model here:
http://physics.weber.edu/schroeder/software/demos/IsingModel.html

25.3 Spatial dimensionality is crucial
For the existence of a phase transition, not only the system size but also the spatial-dimensionality of the system is crucial.

Let us consider a one-dimensional Ising model (Ising chain), whose total energy reads
\[
H = -J \sum_{-\infty < i < +\infty} \sigma_i \sigma_{i+1}.
\] (25.2)
We have ignored the external magnetic field for simplicity. Compare the energies of the following two spin configurations (+ denotes the up spins and − down spins) (Fig. 25.1):

\[
+ + + + + + + + + + + + + + + + + + + + + + + +
\]

\[
+ + + + - - - - - - - - - + + + + + + + + + +
\]

Figure 25.1: Top: completely ordered state of 1-Ising model (+ implies up spins and − down spins); Bottom: Ising chain with a spin-flipped island of size \( L \).

The bottom one has a larger energy than the top by \( 2J \times 2 \) due to the existence of the two mismatching edges. However, this energy difference is independent of the size \( L \) of the island. Therefore, as long as \( T > 0 \) there is a finite chance of making big (= macroscopic) down spin islands amidst the ocean of up spins. If a down spin island becomes large, there is a finite probability for a large lake of up spins on it. This implies that no ordering is possible for \( T > 0 \).

As you can easily guess there is no ordered phase in any one dimensional lattice

\(^{282}\)This argument requires, honestly speaking, details about convergence of the distribution, etc. See the graduate course notes.
system with local interactions for $T > 0$.

### 25.4 There is an ordered phase in 2D

Consider the two-dimensional Ising model with $h = 0$. Imagine there is an ocean of up spins (Fig. 25.2). To make a circular down-spin island of radius $L$, we need $4\pi JL$ more energy than the completely ordered phase.

![Figure 25.2: Peierls’ argument illustrated.](image)

This energy depends on $L$, making the formation of a larger island harder. That is, to destroy a global order we need a macroscopic amount of energy, so for sufficiently low temperatures, the ordered phase cannot be destroyed spontaneously. Of course, small local islands could form, but they never become very large. Hence, we may conclude that an ordered phase could exist at sufficiently low temperatures. Consequently, there must be a phase transition for a two-dimensional system with local short-range interactions. The above argument is known as Peierls’ argument and can be made rigorous.\(^{283}\)

What Peierls actually proved is the following. Prepare a $L \times L$ square lattice, and fix all the edge spins upward (Fig. 25.3). If $L$ becomes large and if $T$ is not low, eventually the probability $P(\sigma_0 = +1)$ of the center spin to be up converges to $1/2$. If this is always true for any $T$, it implies no spin ordering occurs. Peierls demonstrated that $P(\sigma_0 = +1) > 1/2$ at sufficiently low temperatures. An important idea used in the proof is basically the above argument.

We can watch the effect of the outer boundary condition with the aid of the simulation available at [http://www.pha.jhu.edu/~javalab/ising/ising.html](http://www.pha.jhu.edu/~javalab/ising/ising.html) due to J Wasserman (not available now). The standard simulation is with a periodic boundary condition. First you go to a very high temperature, and then drop the temperature...

---

\(^{283}\)There is at least one more crucial factor governing the existence of phase transition. It is the spin dimension: the degree of freedom of each spin. Ising spins cannot point different directions, only up or down (their spin-dimension is 1). However, the true atomic magnets can orient in any direction (their spin dimension is 3). This freedom makes ordering harder. Actually, in 2D space ferromagnetic ordering at $T > 0$ by spins with a spin dimension larger than 1 is impossible.
to $T$. Irrespective of $T$, you will find that the probability of a particular spin to be up and that to be down are equal. Then, choose the boundary condition to be ‘Up.’ Repeat the ‘quenching experiment.’ Now, if $T$ is sufficiently low, as Peierls told us, overwhelmingly you will see up spins throughout the system.

Figure 25.3: All the boundary spins are fixed to be up. What happens to the central spin in the circle in the $L \to \infty$ limit?

Thus we have learned that spatial dimensionality is crucial for the ordering of the phase. Consequently, spatial dimensionality is crucial for the existence of phase transitions for the system with short-range interactions.\textsuperscript{284}

### 25.5 Interaction ‘range’ is also crucial

What happens if the range of interactions is not finite and the intensity of interaction decays sufficiently slowly? Peierls’ argument is still applicable. Obviously, if each spin can interact with all the spins in the system uniformly, an ordered phase is possible even in 1-space. If the coupling constant $J$ decays slower than $1/r^2$, then an order-disorder phase transition is still possible at a finite temperature in one dimensional space.

**Exercise.** Intuitively explain the last statement. \hfill \square

If the interaction is long-ranged, then the system may not behave thermodynamically normally (the fourth law may be violated). However, if interaction is infinitesimal, then thermodynamics may be saved even if the interaction does not decay spatially. As you see below such a system is essentially the system described by the van der Waals equation of state.

### 25.6 Van der Waals model: key ideas

\textsuperscript{284}Here, ‘short range’ implies the interaction vanishes beyond some finite range, or the strength of the interaction decays sufficiently quickly (say, faster than $1/r^{d+1}$).
Van der Waals proposed the following equation of state (van der Walls equation of state):

\[ P = \frac{Nk_B T}{V - Nb} - \frac{aN^2}{V^2}, \tag{25.3} \]

where \( a \) and \( b \) are positive 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 in which molecules can roam from \( V \) to \( V - Nb \); this would modify the ideal gas law to \( P_{HC}(V - Nb) = Nk_B T \). 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.

**25.7 Van der Waals model: ‘derivation’**

Let us ‘derive’ his equation of state from its entropy. We know that the entropy of a classical ideal monatomic gas reads

\[ S = S_0 + Nk_B \log \frac{V}{N} + \frac{3N}{2} k_B \log \frac{E}{N}. \tag{25.4} \]

Notice that \( E \) in this formula is just the kinetic energy \( K \). For the van der Waals model,

1. \( V \to V - Nb \),

2. \( K = E + Nan \), where \( n = N/V \) is the number density, because each particle receives stabilizing interaction from the surrounding ‘mass’ that must be proportional to \( n \), so the real kinetic energy must be the actual \( E \) minus \(-Nan\).

Therefore, the entropy of the van der Waals gas should read

\[ S = S_0 + Nk_B \log \frac{V - Nb}{N} + \frac{3N}{2} k_B \log \frac{E + aN^2/V}{N}. \tag{25.5} \]

This gives

\[ \frac{1}{T} = \frac{\partial S}{\partial E}\bigg|_V = \frac{3}{2} \frac{Nk_B}{E + aN^2/V}, \tag{25.6} \]

and

\[ \frac{P}{T} = \frac{\partial S}{\partial V}\bigg|_E = \frac{Nk_B}{V - Nb} + \frac{3}{2} \frac{Nk_B}{E + aN^2/V} \left( -\frac{aN^2}{V^2} \right). \tag{25.7} \]

Combining these two, we obtain the van der Waals equation of state:

\[ \frac{P}{T} = \frac{\partial S}{\partial V}\bigg|_E = \frac{Nk_B}{V - Nb} - \frac{aN^2}{TV^2}. \tag{25.8} \]
25.8 Liquid-gas phase transition described by the van der Waals model
The most noteworthy feature of the equation is that liquid and gas phases are described by a single equation.

Let us study the general behavior of (25.8). The first term is \( \propto \frac{1}{(V - Nb)} \), so it blows up at \( V = Nb \). It is basically the ideal gas equation translated by \( Nb \) along the \( V \)-axis. The second term becomes very important if \( T \) is small and is \( \propto \frac{1}{V^2} \). Since \( \frac{1}{V^2} \) becomes larger more quickly than \( \frac{1}{(V - Nb)} \) if \( V \) is not too close to \( Nb \), (25.8) becomes non-monotonic as a function of \( V \) for sufficiently low \( T \). That is, as in Fig. 25.4, the \( PV \)-curve wiggles. We know, however, thermodynamically, \( (\partial P/\partial V)_T \) cannot be positive (it is a Le Chatelier’s principle!). This ‘wrong behavior’ must be due to the attractive interactions. Van der Waals guessed that there is a gas-liquid phase transition. Thus, he proposed some ad hoc equilibrium condition to connect the gas and liquid branches of his equation.

25.9 Maxwell’s rule
Maxwell was fascinated by the equation because of its possibility to describe the liquid and the gas phases in a unified fashion, and gave the liquid-gas coexistence condition (Maxwell’s rule, see Fig. 25.4). As will be discussed below, this equation of state reinforced by Maxwell’s rule can be obtained from statistical mechanics. Thus, the van der Waals model is an exact example exhibiting a phase transition.

Maxwell’s rule is motivated by the calculation of the Gibbs free energy \( G \): \( dG = VdP \) (under constant \( T \))\(^{285}\) (see Fig. 25.5).

25.10 Kac potential and van der Waals equation of state
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

\(^{285}\)Since we cannot use thermodynamics where the system is unstable, the demo here (the original) is not a very respectable one. However, the obtained rule can even be thermodynamically justifiable (as explained in the graduate notes).
Figure 25.4: 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), phase separation occurs (see Lecture 28). The liquid-gas coexistence pressure for a given temperature is determined by Maxwell’s rule: the two shaded regions have the same area.

Figure 25.5: Maxwell’s rule is derived from the equality of the Gibbs free energy along the coexisting line (the horizontal line in Fig. 25.4). The figure depicts the needed integral (signed area $\int V dP$) along the curve. If the green area and the red area become identical, the vertical thick dotted line denotes the coexistence condition.

The idea is

$$P = P_{HC} - \frac{1}{2} \varepsilon n^2, \quad (25.9)$$

where $P_{HC}$ is the hard core fluid pressure, $^{286}$ and the subtraction term is the average effect of attractive forces. 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 $^{287}$ However, this equation

$^{286}$For a collection of hard cores there is no gas-liquid transition at any temperature.

$^{287}$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 interaction energy is bounded from below, then the normal thermodynamics is guaranteed.
augmented with Maxwell’s rule is thermodynamically legitimate, and indeed it is the equation of state of the gas interacting with the Kac potential:

$$\phi(r) = \phi_{HC}(r/\sigma) + \gamma^3 \phi_0(\gamma r/\sigma),$$

where $\phi_{HC}(x)$ is the hard core potential: 0 beyond $x = 1$ and $\infty$ for $x \leq 1$, $\sigma$ is the hard core diameter, and $\phi_0$ is an attractive tail.\(^{288}\) The parameter $\gamma$ is a scaling factor; we consider the $\gamma \to 0$ limit (long range but infinitesimal interaction).

### 25.11 1D Kac potential system may be computed exactly

$P_{HC}$ is not exactly obtainable if the spatial dimension is 2, 3, ..., but in 1-space, we can obtain it exactly. Therefore, the 1D Kac model (augmented van der Waals model) is exactly solvable, and exhibits a phase transition.

First, let us use the fact that the ideal gas law can be obtained purely mechanically as Bernoulli demonstrated long long ago. Look at the trajectories of the particles (diameter or length $\sigma$) (Fig. 25.6). Collisions are just exchange of velocities (the momentum conservation). Therefore, if we trace the trajectories of the centers of mass of the particles disregarding the particle sizes, then the trajectories behave just as shown in the right of Fig. 25.6: point masses going through each other ballistically (ideal gas!). Therefore, the equation of state of the hard ‘rods’ must be identical to the ideal gas law in the space not covered by the particles, that is $L - N\sigma$ or, writing $L$ as $V$ the volume,

$$P(V - N\sigma) = Nk_BT.$$  \hspace{1cm} (25.11)

So we have shown that statistical mechanics can describe phase transitions. Now, let us survey how to study the phase diagram for a given substance.

### 25.12 What are the key points to study the phases?

We wish to map out the equilibrium states in the space spanned by, say, $T$, $P$ and other intensive parameters (usually) as we have seen in Fig. 24.1. To study an ordinary geographical map usually we pay attention to the territorial boundaries first.

\(^{288}\)Notice that the second term of (25.10) is chosen so that

$$4\pi \int_0^\infty \gamma^3 \phi_0(\gamma r/\sigma)r^2dr$$

converges in the $\gamma \to 0$ limit, a wide but shallow potential.
This means we must understand phase transitions. As we will learn near the phase transitions (esp., second order phase transitions) fluctuations become so big that any theory ignoring them does not make sense. However, far away from phase transition points, we may often ignore (at least qualitatively) fluctuations and simple theoretical methods often work.

Thus, the study of the phase diagram consists of two pillars: renormalization-group theory that can handle violent fluctuations and mean-field theory that is convenient if we may ignore fluctuations. We will discuss the mean field theory fairly in detail in this course. The basic idea of renormalization group theory and why it is important will be discussed in this course, but we will not have enough time to discuss its technical aspects.

25.13 Magnets, liquids and binary mixtures share some common features
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. 25.7).

Thus, we can discuss the magnetic system as a representative example.

25.14 Ising model in \(d\)-space: a brief review
We already know spatial dimensionality is crucial for the existence/non-existence of phase ordering and consequently phase transitions. Phase ordering is possible be-
cause the order can resist thermal fluctuation. To this end microscopic entities must stand ‘arm in arm.’ The number of entities with which each entity directly interacts (cooperates) crucially depends on the spatial dimensionality. This is the intuitive reason why spatial dimensionality matters.

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 in Lecture 25 by, e.g., the transfer matrix method; the phase transition does not occur for $T > 0$ as we have intuitively seen above (Peierls’ argument).

2-Ising model:
(1) The Onsager solution gives the free energy without magnetic field. There is a phase transition at $T_c > 0$ as we already know.
(2) Below the phase transition temperature $T_c$ there are only two phases corresponding to the up spin phase and the down spin phase, and there is no phase coexistence. That is, up and down phases cannot coexist (see Fig. 25.8).

---

289It is not hard to show that $T = 0$ is a phase transition point for this model.
290due to L. Onsager.
291Independently due to M. Aizenman and Y. Higuchi.
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).

(2) It is known that at sufficiently low temperatures there are phase coexistence.\(^{293}\)

(3) The critical divergences are non-trivial just as in 2-space.

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.\(^{294}\)

---

25.15 Fluctuation can be crucial in $d < 4$

As you have seen above, near the critical point, or the point where ‘strong order’ disappears at last, fluctuation is quite important if the spatial dimensionality is less than 4. Beyond 4D, however, the effect of fluctuations may not be so pathological, and perhaps we may largely ignore fluctuation effects. This observation is relevant to the study of phase transitions as we will see soon.

\(^{292}\)Here, ‘non-trivial’ means that the fluctuation is so large that we cannot use mean-field theory to study the divergent behavior correctly.

\(^{293}\)due to R. L. Dobrushin

\(^{294}\)due to M. Aizenman. 4-Ising model still defies mathematical studies.
Q25-1. Are the following statements correct or incorrect? If incorrect, you must explain why or give a counterexample for the statement. If correct, you must provide a brief supporting argument. However, if you can quote or point out the relevant statements in the lecture notes, you need not state your own. [Hint: all the answers are written somewhere in the lecture notes.]

1. One of the thermodynamic densities must exhibit discontinuity for a phase transition to occur.
2. When a solid phase melts to a liquid phase, the entropy always increases.
3. No 1D system can exhibit phase transition if the interaction range is finite.
4. When a first order phase transition occurs between phase I and II, these two phases can coexist.

Soln.

1. No. This is required only for first order phase transitions.
2. No. There is a counterexample. For $^3$He, crystallization localizes atoms, so the spin-spin coupling due to exchange of particles is reduced, and the spin order that exists in liquid is lost.
3. Yes. This is according to Peierls’ argument or from the PF theorem.
4. No. A counterexample is the 2-Ising model.

Q25-2. When the temperature is raised under constant pressure, phase I changes into phase II through a first order phase transition. Show that the transition from I to II requires absorption of heat (latent heat).

Soln.

This is basically answerable with the aid of le Chatelier’s principle, BUT, $S$ is NOT differentiable at the phase transition point. However, the basis of the stability argument from which le Chatelier’s principle is derived is the convexity of $E$ as a function of $S$, and other extensive quantities. This means when $\Delta S \Delta T > 0$, so $\Delta S > 0$. The latent heat is this times the transition temperature, so it must be positive (absorbs heat).

Q25-3. The van der Waals equation of state reads

$$P = \frac{Nk_B T}{V - Nb} - a \frac{N^2}{V^2}. \quad (25.12)$$

1. Find the critical temperature $T_c$, volume $V_c$ and pressure $P_c$ in terms of $a, b$ and $k_B$. [$T_c$ is the temperature, where the $PV$-curve has an inflection point.]
2. Van der Waals introduced the concept of reduced pressure $p = P/P_c$, reduced temperature $t = T/T_c$, and reduced volume $v = V/V_c$, and showed that $p$ is a universal function of $t$ and $v$ (called the reduced equation of state). Find this relation.
Figure 25.9: The slope is $T$. The straight portion of $E$ is the coexisting phase of 1 and 2. Since $E$ must be convex, the slope does not decrease. Therefore, the low temperature phase must correspond to the smaller-entropy phase.

**Soln.**

(1)

\[ \frac{\partial P}{\partial V}_T = -\frac{Nk_B T_c}{(V_c - Nb)^2} + 2a \frac{N^2}{V_c^3} = 0, \]  
(25.13)

\[ \frac{\partial^2 P}{\partial V^2} \bigg|_T = 2 \frac{Nk_B T_c}{(V_c - Nb)^3} - 6a \frac{N^2}{V_c^4} = 0, \]  
(25.14)

From this we get

\[ \frac{k_B T_c}{(V_c - Nb)^2} = 2a \frac{N}{V_c^3}, \quad \frac{k_B T_c}{(V_c - Nb)^3} = 3a \frac{N}{V_c^4}, \]  
(25.15)

Therefore (taking the ratio of the above equalities), we get

\[ V_c - Nb = 2V_c/3 \Rightarrow V_c = 3Nb. \]  
(25.16)

Therefore,

\[ k_B T_c = \frac{2aN(2Nb)^2}{(3Nb)^3} = \frac{8a}{27b}. \]  
(25.17)

Thus,

\[ P_c = \frac{N}{(3Nb - Nb)} \frac{8a}{27b} - \frac{a}{9b^2} = \frac{a}{27b^2}. \]  
(25.18)

(2) We have only to rewrite the van der Waals equation in terms of $p$, $t$, and $v$.

\[ p = \frac{a}{27b^2} \frac{N8at/27b}{3vNb - Nb} - a \frac{N^2}{(3Nb)^2}. \]  
(25.19)

That is,

\[ p = \frac{27b^2}{a} \frac{N8at/27b}{3vNb - Nb} - \frac{27b^2}{(3v^2)^2} = \frac{8t}{3v - 1} - \frac{3}{v^2}. \]  
(25.20)
26 Why critical phenomena are difficult; mean-field theory

Summary
* How order is lost upon changing $T$ is discussed.
* It is explained why the second order phase transition is difficult to study.
* To understand the phase diagram we need renormalization group and mean-field theory.
* Why large fluctuations near the critical point imply universality is intuitively explained.
* Rushbrooke’s inequality is demonstrated (as a review of thermodynamics).
* The mean field theory is formulated with the aid of conditional expectations.

Key words
correlation length, central limit theorem, renormalization group, (genuine and trivial) universality, mean-field theory, bifurcation.

What you should be able to do
* You must be able to illustrate what happens if you approach a second order phase transition by changing $T$.
* You must be able to set up the mean-field equation. There are many ways, but the formulation in terms of the conditional probability is the most elegant, so memorize the approach.
* You should know qualitative feature of $\tanh x$.
* How to solve (or qualitatively understand) mean-field equation graphically.
* You should clearly recognize the limitations of the mean field approach.
* You should be able to explain why critical fluctuations help universality.

26.1 Summary up to this point
What is the phase transition? If a thermodynamic variable is varied and if a mathematical singularity (loss of continuity, loss of differentiability, etc.; loss of analyticity, in short) in a thermodynamic potential is observed, we say there is a phase transition. That is, a phase transition is characterized by a mathematical singularity.
of a thermodynamic potential.

The existence of a singularity requires a very large system (strictly speaking, an infinitely large system). Thus, to formulate phase transitions mathematically, we need the thermodynamic limit. In the thermodynamic limit extensive quantities are all infinite, so we define thermodynamic densities (extensive quantities per particle). A phase transition may also be characterized by a mathematical singularity in a thermodynamic density. We have actually seen that statistical mechanics can describe phase transitions (2-Ising model, D Kac-model, although rather peculiar, the Bose-Einstein condensation).

Phase transitions are often studied by changing intensive parameters (e.g., temperature and pressure). When two phases coexist, they share the same intensive parameters (called fields). Therefore, a convenient thermodynamic potential is the generalized Gibbs free energy for a given amount of the material \(N\), that is, the (generalized) Gibbs free energy 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 = fields).\(^{296}\) 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.\(^{297}\)

### 26.2 How is order lost in first order phase transition?

Let us consider how an order is lost upon changing temperature.\(^{298}\) In the case of the first order phase transition, as briefly discussed in the preceding Lecture, decrease in order weakens cooperative interactions, accelerating further decrease of order.\(^{299}\) For example, in the case of a liquid crystal in which slender molecules align in the (nematic) liquid crystal phase, increase of temperature randomizes the molecular alignment and make molecular packing harder, causing increase of the volume.

\(^{296}\)\(G\) must be continuous. Why?

\(^{297}\)There is an infinite order phase transition, where \(G\) maintains to be \(C_\infty\) (infinite-times differentiable) but loses analyticity, that is, formal Taylor expansion ceases to converge. This actually happens in 2D XY model. Also such a singularity occurs between the stable and metastable branches of the free energy.

\(^{298}\)Warning: This is NOT really a thermodynamic or equilibrium statistical-mechanical explanation.

\(^{299}\)If a slight local loss of order could induce something like a domino effect of loss of global order, it is likely a first order phase transition. The Lindemann criterion illustrates this point; although the phase transition is determined by the equilibrium condition of the chemical potentials in the ordered and the disordered phases, instability in the ordered phase seems to occur sufficiently close to the equilibrium order-disorder phase transition point in the solid-liquid phase transition case.
Needless to say, this drastically helps randomization of molecular directions, and the liquid crystal-isotropic liquid phase transition is first order.

### 26.3 Typical second order phase transition

As a typical second order phase transition, let us consider the Ising model. Let us approach the phase transition from the high temperature side. Up and down spin islands grow as $T$ is reduced (see Fig. 26.1; the figure is constructed using [http://www.pha.jhu.edu/~javalab/ising/isings.html](http://www.pha.jhu.edu/~javalab/ising/isings.html) (no more available?); the following site is also useful: [http://physics.weber.edu/schroeder/software/demos/IsingModel.html](http://physics.weber.edu/schroeder/software/demos/IsingModel.html)).

![Figure 26.1: Temperature dependence of spin fluctuations. The right-most figure corresponds to the critical point. The upper half is the disordered high temperature phase, and the lower half is the ordered phase. The correlation length increases from left to right.](image)

### 26.4 Correlation length

The typical size $\xi$ of the islands is called the correlation length.\(^{300}\) We see $\xi$ grows and

\(^{300}\langle s(r)s(0) \rangle \sim e^{-|r|/\xi}$ defines $\xi.$
actually it diverges at $T_c$ as $\xi \sim |T - T_c|^{-\nu}$, where $\nu$ is a positive universal constant (one of critical indices). From the low temperature side, in the almost completely ordered phase appear the opposite spins like blinking stars. They become spin-flipped islands, growing bigger and bigger as $T$ increases (i.e., $\xi$ grows). Eventually, these islands coalesce to make a supercontinent.

26.5 Critical fluctuations and slowing down
The patches (islands) of size $\sim \xi$ appear and disappear, so big fluctuations occur near the critical point. This is the critical fluctuation. Since large scale change cannot be completed very quickly, we see increasing space-time scale fluctuations when $T$ approaches $T_c$ from its either side. That is, the dynamics becomes sluggish near the critical point (called critical slowing down). These fluctuations are not only big but also statistically highly correlated: the spins within $\xi$ behave similarly, so as we approach $T_c$ spin fluctuations become increasingly statistically correlated. Thus, even on the scale we can observe the system is not simply macroscopic, governed by the law of large numbers.

If such critical fluctuations occur in fluid, we see critical opalescence, which Einstein wished to understand (and created the thermodynamic fluctuation theory we have already learned in Section 19, esp., 19.3):

http://www.youtube.com/watch?v=cSliO89x7UU

26.6 Scaling invariance of critical fluctuation
Critical fluctuations have a very special property of scaling invariance: from however large distance you observe the critical fluctuations, they look the same as you observe them from, say, 1m away. This is exhibited by the following two You Tube movies due to D. Ashton, excellent and wonderful:

http://www.youtube.com/watch?v=lQxD1PinDbs
http://www.youtube.com/watch?v=MxRddFrEnPc.
We will come back to the second movie soon.

26.7 Divergence of susceptibilities
We know the fluctuation-response relation: generally,

$$\left. \frac{\partial X}{\partial x} \right|_{y, \ldots} = \beta \langle \delta X^2 \rangle. \tag{26.1}$$
For example, the (isothermal) magnetic susceptibility $\chi$ is directly related to the variance of magnetization:

$$\chi_T = \beta(\delta M^2).$$  \hspace{1cm} (26.2)

We just learned the big critical fluctuation, so we can expect that the susceptibilities become very large near the critical point.

![Figure 26.2: Schematic illustrations of singular behaviors near the critical point.](image)

### 26.8 Critical indices

Empirically the susceptibility diverges as (for $h = 0$, without magnetic field) (see Fig. 26.2)

$$\chi \sim |T - T_c|^{-\gamma} = |\tau|^{-\gamma} \quad (h = 0),$$

where $\tau = (T - T_c)/T_c$.

We cannot expect smooth increase of the magnetization $m$ from zero below $T_c$: \[301\]

$$m \sim (-\tau)^\beta \quad (h = 0, \tau < 0).$$  \hspace{1cm} (26.4)

The divergence of energy (or entropy) fluctuation causes the divergence of specific heat as

$$C_B \sim |\tau|^{-\alpha} \quad (h = 0).$$  \hspace{1cm} (26.5)

$\alpha$, $\beta$, $\gamma$, are positive numbers and are called critical indices. Representative values can be found in the table below. They are universal numbers. For example, for any fluid or binary mixture, or magnets (with an easy axis = Ising magnets) these numbers are common. They are determined by the nature of our world, not material-scientific trivial or fetish facts.

---

\[301\] 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.
### Ising critical exponents.

<table>
<thead>
<tr>
<th>Exponents</th>
<th>2-space</th>
<th>3-space</th>
<th>(d(\geq 4))-space</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>0 (log)</td>
<td>0.11</td>
<td>0 (jump)</td>
</tr>
<tr>
<td>(\beta)</td>
<td>1/8</td>
<td>0.325</td>
<td>1/2</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>1.75</td>
<td>1.24</td>
<td>1</td>
</tr>
<tr>
<td>(\delta)</td>
<td>15</td>
<td>4.8</td>
<td>3</td>
</tr>
<tr>
<td>(\nu)</td>
<td>1</td>
<td>0.63</td>
<td>1/2</td>
</tr>
</tbody>
</table>

#### 26.9 Critical index (in)equalities

It was empirically noted that several relations hold among these indices such as

\[
\alpha + 2\beta + \gamma = 2. \tag{26.6}
\]

Thermodynamically, we can prove

\[
\alpha + 2\beta + \gamma \geq 2. \tag{26.7}
\]

Demonstration of this inequality gives us an excellent opportunity to review elementary thermodynamics.

#### 26.10 Proof of index inequality (Rushbrooke inequality)

Even around the critical point the system does not become thermodynamically unstable. Therefore, inequalities required by the thermodynamic stability conditions remain valid. For example, thermodynamic stability of a magnet implies (the Gibbs relation in this case is \(dE = TdS + BdM\))

\[
\frac{\partial(S, M)}{\partial(T, B)} \geq 0. \tag{26.8}
\]

This inequality can be written explicitly, by expanding the determinant defining the Jacobian, as

\[
\frac{\partial S}{\partial T} \bigg|_B \frac{\partial M}{\partial B} \bigg|_T \geq \frac{\partial S}{\partial B} \bigg|_T \frac{\partial M}{\partial T} \bigg|_B = \left( \frac{\partial M}{\partial T} \right)_B^2, \tag{26.9}
\]

where a Maxwell’s relation

\[
\frac{\partial(S, T)}{\partial(B, M)} = 1 \tag{26.10}
\]

---

\(^{302}\)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.
has been used to obtain the second equality. This implies
\[
\frac{1}{T} C_B \chi \geq \left. \frac{\partial M}{\partial T} \right|_B^2.
\] (26.11)

Introducing the definitions of the critical exponents given above, we obtain
\[
|\tau|^{-\alpha} |\tau|^{-\gamma} \geq |\tau|^{2(\beta - 1)}.
\] (26.12)

Here, we have ignored all the finite coefficients near the critical point (such as \(T^{-1}\)).

(26.12) implies that
\[
|\tau|^{-\alpha} |\tau|^{-\gamma} \geq |\tau|^{2(\beta - 1)}.
\] (26.13)

is required for \(\tau \to 0\). Therefore, the quantity in the parentheses must be nonnegative:
\[
\alpha + 2\beta + \gamma \geq 2.
\] (26.14)

This is called \textit{Rushbrooke’s inequality}.

26.11 Fluctuation and universality

Watch Dr Ashton’s movie again. Large scale fluctuations seems to dominate the scene, and the phenomenon looks the same from however far away you observe it. Thus, we could guess that microscopic details should not be very important. That is, we expect the \textit{universality}: microscopic details do not matter for salient features of the phenomena (in this case critical phenomena). Indeed, the critical indices for fluids, binary mixtures, and 3D Ising magnets are identical (cf. Fig. 25.7). The point is impressively illustrated by Dr Ashton.

26.12 Universality, nontrivial and trivial

A typical and good example of universality is the osmotic pressure of polymer solutions. Irrespective of the polymers and solvents, as long as the polymers are long enough, and the solvent dissolves polymers well, Fig. 26.3 is quantitative:

Here is one parameter we cannot theoretically compute, i.e., the proportionality constant between \(X\) and \(c\). However, notice that except for this one parameter, there is no freedom left: even if you change, solvents, polymers, or temperature, Fig. 26.3 is quantitatively correct. That is, an (uncountably) infinite dimensional space of

\[303\] We have assumed that the critical point is not zero; The 1-Ising model has \(T_c = 0\), but this is a pathological example.
polymer solutions is subsumed to a 1-dimensional (one-parameter) world.

You might say we already know examples of universality. For example, we know

\[ PV = \frac{2}{3} E \]  

(26.15)

for any ideal gas irrespective of statistics: if the spatial dimensionality is 3, and the dispersion relation is \( \varepsilon \propto p^2 \), this is true. Or we know \( PV = E/3 \) for phonons and photons in 3-space. This is quite universal. However, it is due to the universality (common quantitative feature) of the elementary entities (without interaction!) making up the system. In this sense, universality is unsurprising, and trivial. Furthermore, if you modify a system a bit with interactions, the \( PV \) relation changes sensitively. That is, infinite ways of changing the system result in infinitely different modifications of the resulting equation of state.

In contradistinction, the universality discussed in the preceding paragraph is obviously not due to some common quantitative features at the microscopic level. Of course, the system must be a polymer system, but we use only two features: a polymer is very long, and cannot cross itself; nothing quantitative is required. In this case, there are infinitely many ways to modify polymer-solvent systems (by simply changing polymers and solvents), but only one parameter is modified in consequence (i.e., only \( X/c \) is modified) as emphasized above. That is, this universal feature is quite stable against materialistic perturbations. Thus, the above mentioned universality is deserved to be called the *genuine universality*.

Elementary statistical mechanics is boring because it emphasizes trivial universality only.
26.13 Study of phase diagram
The purpose of statistical mechanics is to understand various macroscopic properties of a given equilibrium system. We know this is equivalent to mapping out its thermodynamic space. Just as the ordinary maps we are interested in two things: where the borders are and what each territory looks like. The former corresponds to the study of phase transitions and the latter to characterized each phase. As we have already seen in Fig. 24.2 the phase diagram on the thermodynamic space is complete, but usually we are contented with the diagram on, say, the $PT$-space such as Fig. 25.7. Let us look at it again:

![Phase Diagram](image)

Figure 26.4: The shaded area is dominated by fluctuations; if you go away from it along the arrows, the correlation length diminishes and mean field theory becomes respectable.

To understand boundaries (phase transitions) if you are close to the critical point or the second order phase transition point, you must deal with fluctuations and high correlation (statistical dependence). There, many things would be universal as we have seen briefly. If away from the critical point, the phase transitions become first order. The phase transition locations are determined by the equality of the chemical potentials of the phases between which the transition occurs. Thus, many small details including small fluctuations that sensitively modify the chemical potentials really matter. If you are away from the critical point (and phase transitions in general) along the arrows in the figure, then the correlation length reduces and fluctuation becomes less significant (i.e., microscopic entities are more statistically independent).

At this juncture watch Ashton’s ‘zooming out’ movie again. Since there is no qualitative difference along the arrows, to understand the general feature of various phases, we can study the regions where the correlation length is small and fluctuations are small (i.e., the zoomed-out states). Then, at least qualitatively what happens at the first order phase transition can also be understood. For example, to understand ice we could study sufficiently low temperature perfect crystal ice, and
to understand vapor, we can study very high temperature gas (almost an ideal gas).

Thus, the study of the phase diagram consists of two strategies: to roughly understand the nature of phases we may largely ignore fluctuations and to understand phase boundaries we need a means to cope with fluctuations. The former is represented by the mean field theory we will discuss below and the latter by renormalization group theory. We cannot go into the latter in this course, but in the following its crucial mathematical core will be explained.

### 26.14 Central limit theorem

We have realized that near the second order phase transition/critical point, fluctuation becomes large, and also strongly correlated. How can we handle such a strongly correlated fluctuating system? We wish to know the distribution of fluctuations on the mesoscopic scale (because the correlation lengths are on the mesoscopic scale). If we understand the mesoscopic fluctuation statistics, we should be able to compute macroscopic observables.

How can we study the mesoscopic fluctuation statistics? We encounter the third pillar of probability theory: central limit theorem (CLT); recall that the law of large numbers and large deviation theory are the two of the three pillars we have already utilized.

The law of large numbers may be understood as the statement that the density distribution function of \(\frac{1}{N} \sum_{i=1}^{N} X_i\) is concentrated to a single point \(m = \langle X_1 \rangle\), if \(X_i\) are iid. Here comes the other refinement of LLN: the central limit theorem. Instead of \(N\) if we divide the partial sum \(S_N = \sum X_i\) with a smaller quantity, say \(N^{3/4}\), which must be chosen just right, the density distribution of \(S_N/N^{3/4}\) might converge to a nice function. In the case of iid stochastic variables, we know the size of the fluctuation of \(S_N\) is \(\sqrt{N}\), so perhaps \(\sqrt{N}\) is the right factor.

If \(X_i\) are iid\(^{305}\) with the distribution not too broad, this guess is correct. We have the central limit theorem:

**Central Limit Theorem**

If \(X_i\) are assumed to be iid with a finite variance, the density distribution function of \((S_N - Nm)/\sqrt{N}\) converges to a Gaussian function \(N(0, V)\), where

\[
m = \langle X_1 \rangle \quad \text{and} \quad V = \langle (X_1 - m)^2 \rangle.
\]

This allows us to calculate expectation values of quantities dependent on fluctuations.\(^{306}\)

---

\(^{304}\)Recall \(S_N = N\langle X \rangle + O[\sqrt{N}]\).

\(^{305}\)The CLT in this form holds more generally even if \(X_i\)'s are not iid, but are correlated. If the correlation decays sufficiently quickly or if \(X_i\) make a Markov process, the CLT with \(\sqrt{N}\) holds. However, the variables we are interested in here behave much more strongly correlated, and the factor \(\sqrt{N}\) is not appropriate.

\(^{306}\)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 systems with finite variances. 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. Fortunately (or unfortunately), for iid with finite variance they give the same conclusion about
26.15 We need much more general CLT to understand critical phenomena
Unfortunately, the system we are interested in are strongly interacting systems, so \( X_i \) are not at all statistically independent, but strongly correlated rather globally. Therefore, to understand second-order phase transition, we need a vast generalization of the usual CLT to the case with strong correlations. This extension is the renormalization group theory. Let \( X_i \) be the \( i \)th spin. Then, \( M = \sum X_i \) is the magnetization. Let us consider this in the paramagnetic phase (i.e., \( M = 0 \) on the average). We are interested in its fluctuation. Due to strong correlation, now we cannot choose \( y = 1/2 \) in \( M/N^y \) to have a nice distribution function in the thermodynamic limit; we must choose \( y \) in a highly nontrivial fashion, which is related to the critical indices.\(^{307}\)

26.16 Misunderstanding of CLT
In elementary thermal physics courses, CLT is invoked to explain intuitively why the existence of many particles makes macroobservables almost deterministic. However, as clearly noted in this course, the law of large numbers is the fundamental reason. If you know how to prove CLT, you will agree that CLT is much more sophisticated and restricted. Usual use of CLT to explain why statistics works is like showing continuity by showing differentiability. This abuse of math is likely to be caused by misunderstanding of CLT. The main claim of CLT is, after scaling, the nontrivial distribution emerges that is system-size (sample-size) independent. This is renormalization group!

We need only LLD. Every instructor should understand this.

26.17 Mean field Idea
Sufficiently away from critical points/second order phase transition points, the equilibrium average of a function of several spins \( f(s_0, s_1, \cdots, s_n) \) may be computed through separately averaging 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

\[
\langle f(s_0, s_1, \cdots, s_n) \rangle \simeq f(\langle s_0 \rangle, \langle s_1 \rangle, \cdots, \langle s_n \rangle). \tag{26.16}
\]

This is the basic idea of the mean field approach. Here, let us proceed slightly more systematically.

26.18 Quantitative formulation of mean field theory: fundamental equation
Let us look at an elementary identity of probability theory. 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 sample space \( \Omega \)), then (Fig. fluctuations.

\(^{307}\)Beyond this point, you can go to my graduate school lecture notes (Chapter 5) or “Informal Notes on Renormalization and Phase Transitions” (both downloadable). A more elementary explanation may be found in The Nonlinear World (Springer 2012) (Chapter 3).
26.5)

\[ E(E(A|B_i)) = \sum_i P(B_i) E(A|B_i) = E(A). \]  

(26.17)

That is, the average of a conditional expectations over all the conditions is equal to the unconditional average.

\[
\Omega
\]

Figure 26.5: (26.17) is illustrated. Suppose \( A \) is something (shade in the figure) distributed on the sample space \( \Omega \) which is partitioned into \( B_1, \cdots, B_5 \). In each partition we can define the average on it \( E(A|B_i) \). If the probability for event \( B_i \) to happen is \( P(B_i) \), the average of \( A \) over \( \Omega \) is given by (26.17).

Let us choose as \( B_i \) a particular configuration ‘\( i \)’ \( \{s_1, \cdots, s_{2d}\} \) of all the spins interacting with the ‘central spin’ \( s_0 \) on a \( d \)-cubic lattice (Fig. 26.6). Notice that if \( s_1, \cdots, s_{2d} \) are fixed, the central \( s_0 \), which is interacting only with these neighboring spins, is totally decoupled from the rest of the world.

\[
E(s_0|s_1, \cdots, s_{2d}) = \sum_{s_0} s_0 e^{\beta J s_0 (s_1 + \cdots + s_{2d}) + \beta h s_0} \sum_{s_0} e^{\beta J s_0 (s_1 + \cdots + s_{2d}) + \beta h s_0} = \tanh[\beta h + \beta J (s_1 + \cdots + s_{2d})].
\]

(26.18)

Because \( E(s_0) = E(E(s_0|s_1, \cdots, s_{2d})) \), we obtain

\[ \langle s_0 \rangle = \langle \tanh [\beta h + \beta J (s_1 + \cdots + s_{2d})] \rangle. \]  

(26.19)
This is an *exact* relation into which we may introduce various approximations to construct mean field approaches.

### 26.19 The crudest version of the mean-field theory

Now, to compute the RHS of (26.19), we must introduce some approximation. The most popular (and simple-minded) version is (26.16) or more concretely for the present example:

\[
\langle \tanh[\beta h + \beta J(s_1 + \cdots + s_{2d})] \rangle \simeq \tanh[\beta h + \beta J\langle s_1 + \cdots + s_{2d} \rangle].
\]  
(26.20)

Therefore, for \( m = \langle s_0 \rangle \), we obtain a closed equation (consistency equation)

\[
m = \tanh[\beta(2dJm + h)].
\]  
(26.21)

\( 2dJm + h \) may be understood as an effective magnetic 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.

### 26.20 How to solve the consistency equation

Let \( 2d\beta Jm = x \). (26.21) reads

\[
x = 2d\beta J \tanh(x + \beta h).
\]  
(26.22)

For simplicity, let us assume \( h = 0 \). We have to solve

\[
x = 2d\beta J \tanh x.
\]  
(26.23)

This may be graphically solved (Fig. 26.7).

The bifurcation\(^{308}\) 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 \( \beta = 1/2 \)).

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 \( m \neq 0 \) solutions have

---

\(^{308}\) A phenomenon that the solution (or the solution set) changes its character is called *bifurcation*. There are many types, and the one we encounter here 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.
Figure 26.7: The solution to (26.23) may be obtained graphically.

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; if the state is a stable equilibrium, small deviation from the state will decay. The stability of \( m \neq 0 \) state is obvious.

Figure 26.8: The stability of the solution to (26.23) may also be understood graphically.

### 26.21 How reliable is the mean field theory?

We have introduced the idea of mean field theory to study the system thermodynamics 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 we just discussed) 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.

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 and if a mean field theory still tells us that there is no ordering, this assertion is likely to be true. For the ferromagnetic Ising model this expectation has been vindicated. The same idea tells us that the mean field critical temperature should be the upper bound of the true critical temperature: $T_c \leq T_{c\text{, mean}}$.

26.22 Use of mean-field approach: practical guide

Generally speaking, mean-field approaches may be relied upon, if the fluctuation effect is not decisive:

(1) If the spatial dimensionality is sufficiently high, then ‘spins’ gang up against thermal fluctuations;

(2) If the first order phase transition is with a big ‘jump,’ then fluctuations may not easily be able to fill the gap that must be jumped.

Thus, these cases are (often) amenable to mean-field approaches (at least qualitatively).

Perhaps, practically, we may summarize the use of mean field method as follows: we should not swallow the results of the method uncritically (especially as to the phase transitions), but since the method is easy to use in many cases, it is worth trying first.
27 Improving mean field and transfer matrix

Summary
* Mean field theory can be somewhat improved, if the constraints imposed on the spin variables are honestly taken into account.
* Transfer matrix technique is outlined.

Key words
mean-field theory, transfer matrix, Perron-Frobenius’ theorem, Perron-Frobenius eigenvalue

What you should be able to do
* Remember that we should respect the algebraic structure inherent in the system.
* You must be able to set up the self-consistency equation for mean field approaches.
* You must practically be able to set up transfer matrices for 1D finite range models.
* Memorize the Perron-Frobenius theorem.

27.1 Naive mean field approach: a review
Let us review the simplest mean-field approach in detail: the 1D Ising chain whose Hamiltonian is

\[ H = -J \sum_i s_is_{i+1} - h \sum_i s_i. \]  

(27.1)

We have derived the fundamental equality, which reads for the present case as

\[ \langle s_0 \rangle = \langle \tanh[\beta J(s_{-1} + s_1) + \beta h] \rangle. \]  

(27.2)

The naivest mean-field approach is to use the following type of approximation: \( \langle f(x) \rangle \simeq f(\langle x \rangle) \). Since we may assume that the equilibrium single phase is translationally symmetric, \( \langle s_0 \rangle = \langle s_{\pm 1} \rangle = m \). Then, with the above mentioned approximation (27.2) becomes

\[ m = \tanh[2\beta Jm + \beta h]. \]  

(27.3)

(27.3) reads

\[ 2\beta Jm + \beta h = 2\beta J \tanh[2\beta Jm + \beta h] + \beta h \]  

(27.4)

or

\[ x = 2\beta J \tanh x + \beta h. \]  

(27.5)
Here, we will not discuss \( h \), so let us set \( h = 0 \). Then,

\[
x = 2\beta J \tanh x.
\]

(27.6)

Thus, \( 2\beta J = 1 \) determines the critical point: \( T_c = 2J/k_B \).

Everyone knows this is wrong. Intuitively, ordering is hindered by fluctuations, so, if you ignore the effect of fluctuations, then ordered phases tend to be stable at higher temperatures. Therefore, the order-disorder transition point estimated by a mean-field approach tends to be an overestimation at best; the predicted phase transition may not even exist as in this case.

### 27.2 Improving mean field approach

There is, however, a room to improve the mean field theory. We know \( s^2 = 1 \). Using this, we can handle fluctuations in a better way. Let us expand (27.2). Notice that \( \tanh x \) can be expanded into an odd power series:

\[
\tanh x = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots.
\]

(27.7)

In our case \( h = 0 \), so we need odd powers of \( s_{-1} + s_1 \). For example,

\[
(s_{-1} + s_1)^3 = s_{-1}^3 + 3s_{-1}^2s_1 + 3s_{-1}s_1^2 + s_1^3 = s_{-1} + 3s_1 + 3s_{-1} + s_1 = 4(s_{-1} + s_1).
\]

(27.8)

Analogously, any odd power of \( s_{-1} + s_1 \) is proportional to \( s_{-1} + s_1 \). Therefore, we must have the following identity if \( s_i \) takes only \( \pm 1 \) with an appropriate constant \( A \):

\[
\tanh[\beta J(s_{-1} + s_1)] = A(s_{-1} + s_1).
\]

(27.9)

\( A \) is determined by substituting \( \pm 1 \) to the spins:

\[
\tanh(2\beta J) = 2A.
\]

(27.10)

Therefore, (27.2) reads (with \( h = 0 \))

\[
\langle s_0 \rangle = \frac{\tanh 2\beta J}{2}\langle s_{-1} + s_1 \rangle.
\]

(27.11)

That is, we have obtained an exact relation:

\[
m = m \tanh(2\beta J).
\]

(27.12)
We know $|\tanh(2\beta J)| < 1$, so unless $T = 0$, $m = 0$. Therefore, there is no phase transition for $T > 0$. Mean field theories may not be that bad!\footnote{However, this seems to be a very lucky case; we have an exact formula! Generally speaking, you must not trust mean field theoretical results too much as to the phase transitions.}

At this juncture, reread $26.22$, especially, when mean-field approaches may be OK.

### 27.3 Transfer matrix method

To conclude the second-order phase transition, let us discuss one general method to get the free energy exactly. Let us consider a 1D-Ising model with the Hamiltonian\footnote{Strictly speaking, the term proportional to $J$ is the Hamiltonian of the system itself, and the term proportional to $h$ is the potential energy the system has between the system making the magnetic field $h$.} given by

$$
H = -J \sum_{i=1}^{N-1} s_i s_{i+1} - h \sum_{i=1}^{N} s_i. 
$$

Let us define the partition function $Z_N(\uparrow)$ for the length $N$ spin chain with the $N$th spin up:

$$
Z_N(\uparrow) = \sum_{\{s_n\}_{n=1}^{N-1}} e^{\beta[J+(1)s_N-1+h(1)]} e^{\beta[J s_{N-1}s_{N-2}+hs_{N-1}]} \cdots e^{\beta[J s_2 s_1+h s_2]} e^{\beta h s_1}. 
$$

We can analogously define the partition function $Z_N(\downarrow)$ for the length $N$ spin chain with the $N$th spin down. In terms of these, we can make $Z_{N+1}(\pm)$ as

$$
Z_{N+1}(\pm) = \sum_{s=\pm1} e^{\pm\beta[J s+h]} Z_N(s). 
$$

Therefore, if we introduce the vector

$$
Z_N = \begin{pmatrix} Z_N(\uparrow) \\ Z_N(\downarrow) \end{pmatrix},
$$

we can write

$$
Z_{N+1} = T Z_N,
$$

(27.17)
where $T$, called the transfer matrix,\[^{311}\] is defined as

$$T = \text{Matr}(e^{\beta [J_{ss'} + hs']}) = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J + \beta h} \\ e^{-\beta J - \beta h} & e^{\beta J - \beta h} \end{pmatrix}. \hspace{1cm} (27.18)$$

Notice that

$$Z_N = (1, 1) Z_N. \hspace{1cm} (27.19)$$

Repeated use of the recursion (27.17) results in

$$Z_N = T_N \begin{pmatrix} e^{\beta h} \\ e^{-\beta h} \end{pmatrix}. \hspace{1cm} (27.20)$$

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, $Z_N = T_1 T_N$.

### 27.4 How to compute the product of matrices

The easiest method to compute (27.20) is to use an orthogonal transformation (or more generally unitary transformation) to convert $T$ into a diagonal form.\[^{312}\]

$$T = U^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U, \hspace{1cm} (27.21)$$

where $\lambda_1$ and $\lambda_2$ are eigenvalues of $T$, and $U$ is the orthogonal transformation needed to diagonalize $T$. Introducing (27.21) into (27.20), we obtain

$$Z_N = U^{-1} \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} U \begin{pmatrix} e^{\beta h} \\ e^{-\beta h} \end{pmatrix}. \hspace{1cm} (27.22)$$

Therefore, we finally have the following structure:

$$Z_N = a\lambda_1^N + b\lambda_2^N. \hspace{1cm} (27.23)$$

---

\[^{311}\] *The origin of the transfer matrix method* \hspace{1cm} 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.]

\[^{312}\] If impossible, in a Jordan normal form. A necessary and sufficient condition for a matrix $T$ to be diagonalizable is that it is normal: $T^* T = T T^*$. If all the eigenvalues are distinct, the matrix is normal. In the present case, the eigenvalues are distinct, so the matrix is diagonalizable.
where \(a\) and \(b\) are nonzero 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

\[
f = -k_B T \log \lambda_1.
\]  
(27.24)

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.

### 27.5 Why there is no phase transition in 1-space

Let us discuss why there is no phase transition for \(T > 0\) in 1D finite-range interaction systems from the transfer matrix point of view. The free energy could exhibit singularity if \(Z \leq 0\), but this does not happen, because \(Z\) is a sum of positive terms. As long as \(T > 0\), \(K\) is finite, so all the elements of the transfer matrix are without any singularity as a function of \(T\)'s (and \(h\)), and eigenvalues are algebraic functions\(^3\) of the entire functions of \(T\) and \(h\). Therefore, as long as eigenvalues are finite, their singularities are branch points.\(^4\) 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]\(^5\)**

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 \(\lambda\),\(^6\)

(ii) the elements of the eigenvector belonging to \(\lambda\) may be chosen all positive. \(\square\)

This special real eigenvalue giving the spectral radius is called the *Perron-Frobenius*\(^7\)**

---

\(^3\)The elements of the transfer matrix are entire functions; a function that is holomorphic (= no singularities) except at infinity is called an *entire function*.

\(^4\)An algebraic function of \(\{x_i\}\) is a function that can satisfy a polynomial whose coefficients are polynomials of \(\{x_i\}\). An entire function is a function whose singularity is only at infinity.

\(^5\)For example, consider \(\sqrt{z}\). This is real if \(z > 0\) and have two values, but for \(z = 0\) there is only one value. In this case, \(z = 0\) is an example of the branch point. Take \(x^2 - 2zx + 1 = 0\). The roots are algebraic functions of \(z\): \(x = z \pm \sqrt{z^2 - 1}\). Therefore, \(z = 1\) is a branch point for \(x\). Notice that \(dx/dz = 1 \pm z/\sqrt{z^2 - 1}\), so at the branch point, the derivative ceases to exist (for this example).

\(^6\)For a succinct proof, see the P504 course note.

\(^7\)That is, \(\lambda\) gives the spectral radius of \(A\).
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.

27.6 Onsager obtained the exact free energy of 2-Ising model

Onsager used the transfer matrix method to evaluate the partition function of the 2-Ising model on the square lattice exactly. 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.

Exact solutions are very useful of course, but the reasons for the possibility of exact solutions could be rather unimportant peculiarities from the physics point of view. Therefore, it is not productive to rely on exact solutions to construct general theories.

---

318weis 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 an 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.

319Onsager’s much greater contribution to statistical physics is his contribution to nonequilibrium theory, which we have glanced at already. This point seems often ignored as we read explicitly in S. G. Brush, Statistical Physics and the Atomic Theory of Matter, from Boyle and Newton to Landau and Onsager (Princeton UP, 1983).
Q27-1. Consider Ising spins on the diamond lattice (without an external magnetic field $h$). The interactions of the spins are restricted to the nearest neighbor pairs.

1. Write down the fundamental equation for this system corresponding to 
\[ \langle s_0 \rangle = \langle \tanh [\beta J(s_1 + \cdots + s_{2d})] \rangle \] 
(27.25) on the $d$-cubic lattice.

2. What is the critical point, if you use the simplest mean-field theory that uses the approximation $\langle \tan(\cdots) \rangle = \tan(\langle \cdots \rangle)$?

3. This is a three-dimensional system, so there is definitely a positive critical temperature $T_c$. What can you say about this true $T_c$ from your result in (2)?

**Soln.**

1. This is quite the same as is explained in the notes. On the diamond lattice one spin has only 4 nbh spins, so 
\[ \langle s_0 \rangle = \langle \tanh [\beta J(s_1 + \cdots + s_4)] \rangle, \] 
(27.26) where $s_1, \cdots, s_4$ are the spins connected to $s_0$ with the C-C bonds.

2. The naivest approach gives 
\[ \langle s_0 \rangle = \tanh [\beta J \langle s_1 + \cdots + s_4 \rangle]. \] 
(27.27)

Expecting the translationally symmetric magnetization, 
\[ m = \tanh 4\beta J m \] 
(27.28) is the crudest mean-field equation. This means $x = 4\beta J \tanh x$, so $T_c = 4J/k_B$.

3. Thermal fluctuation tends to be against ordering, so theories ignoring fluctuations overestimate the ordering effect, pushing the critical point up. Thus, we may conclude that the critical temperature of the Ising model on a diamond lattice must not be higher than $4J/k_B$. Actually, we may guess the true $T_c$ is far less than this. ($T_c = 2.7J/k_B^{320}$)

Q27-2. We have derived in 26-1 the fundamental equation for the starting point of the mean field approaches for the diamond lattice as 
\[ \langle s_0 \rangle = \langle \tanh [\beta J(s_1 + \cdots + s_4)] \rangle. \] 
(27.29)

Let us be better than 26-1. We wish to exploit the fact that $s^2 = 1$.

1. Expanding tanh in a power series, show that 
\[ \tanh[\beta J(s_1 + \cdots + s_k)] = A(s_1 + s_2 + s_3 + s_4) + B(s_1 s_2 s_3 + s_1 s_3 s_4 + s_2 s_3 s_4 + s_1 s_2 s_4). \] 
(27.30)

---

That is, any odd power of \((s_1 + s_2 + s_3 + s_4)\) is written as a sum of \((s_1 + s_2 + s_3 + s_4)\) and \((s_1s_2s_3 + s_1s_3s_4 + s_2s_3s_4 + s_1s_2s_4)\).

(2) Determine \(A\) and \(B\) by setting \(s = \pm 1\) so that (27.30) holds, or show that

\[
A = \frac{1}{8}(\tanh 4\beta J + 2 \tanh 2\beta J). \tag{27.31}
\]

(3) Now, introducing (27.30) into (27.29), we get the following equation

\[
\langle s_0 \rangle = A \langle s_1 + s_2 + s_3 + s_4 \rangle + B \langle s_1s_2s_3 + s_1s_3s_4 + s_2s_3s_4 + s_1s_2s_4 \rangle. \tag{27.32}
\]

\(\langle s_0 \rangle = \langle s_1 \rangle = \cdots = m\) is the magnetization per spin, so (27.32) reads

\[
m = 4Am + 4B\langle s_1s_2s_3 \rangle. \tag{27.33}
\]

Notice that up to this point there is NO APPROXIMATION, but, unfortunately, we cannot solve (27.33). Now, let us introduce the approximation

\[
\langle s_1s_2s_3 \rangle = m^3. \tag{27.34}
\]

Then, our ‘approximate’ mean field equation is

\[
m = 4Am + 4Bm^3. \tag{27.35}
\]

What is the condition that determines the phase transition? [Hint. At what value of \(A\) is there a bifurcation\(^{321}\)?]

**Soln.**

(1) Checking first 2 or three terms in the expansion of \(\tanh\) is practically all right.

However, if we wish to ‘prove’ (27.30), we can proceed as follows. Generally we have an odd power \((s_1 + s_2 + s_3 + s_4)^m\), where \(m\) is an odd positive integer. If we expand this, (e.g., multinomial theorem) we obtain for \(a + b + c + d = m\) \((a, \cdots, d\) are nonnegative integers)

\[
s_1^a s_2^b s_3^c s_4^d.
\]

There is a perfect permutation symmetry among \(s_1, \cdots, s_4\), so we have only to consider the types of terms as follows [Do not honestly expand \(\tanh\)]

\[
s_1^m = s_1,
\]
\[
s_1^{m-1} s_2 = s_2,
\]
\[
s_1^{m-2} s_2^2 = s_1,
\]
\[
s_1^{m-2} s_2 s_3 = s_1s_2s_3.
\]
\[
s_1^{m-3} s_2^3 = s_2,
\]
\[
s_1^{m-3} s_2^2 s_3 = s_3
\]
\[
s_1^{m-3} s_2 s_3 s_4 = s_2 s_3 s_4
\]

\(^{321}\)Recall ‘bifurcation’ implies the change of number of (real) roots.
Therefore, summing all these terms, we must have (27.30).
(2) For all \( s \) being +1 case:

\[
\tanh(4\beta J) = 4A + 4B.
\]

For one \(-1\)

\[
\tanh(2\beta J) = 2A - 2B.
\]

Other possibilities do not give any new relation. From these, we get

\[
A = (1/8)(\tanh 4\beta J + 2 \tanh 2\beta J), \quad B = (1/8)(\tanh 4\beta J - 2 \tanh 2\beta J).
\]

(3) We must solve \( m = 4Am + 4Bm^3 \). One way is to follow the lecture notes, i.e., a graphical method. This tells us that when the slope of \( 4Am + 4Bm^3 \) at \( m = 0 \) is 1, bifurcation occurs. Hence, \( 4A = 1 \) or

\[
\tanh 4\beta J + 2 \tanh 2\beta J = 2. \tag{27.36}
\]

You need not solve this, but notice that the \( T_c \) obtained from this must be smaller than that obtained from \( 4\beta J = 1 \) (due to a better approximation).

**Q27-3.** At each lattice site of 1D lattice is a particle which can take the ground state and excited state. Only nearest neighbor excited state can interact and the excitation energy required is \( \varepsilon \ (> 0) \). The Hamiltonian may be written as

\[
\mathcal{H} = -J \sum_i \sigma_i \sigma_{i+1} + \varepsilon \sum_i \sigma_i, \tag{27.37}
\]

where \( \sigma_i = 0 \) denotes the ground state, and \( \sigma_i = 1 \) the excited state. Find the free energy per particle (i.e., write down the transfer matrix and compute its eigenvalues).

**Soln.**

The transfer matrix can be made as follows.

\[
T = \begin{pmatrix}
0 & 1 \\
1 & e^{-\beta \varepsilon}
\end{pmatrix}
\]

\[
e^{-\beta J-\beta \varepsilon}
\]

Therefore, the characteristic equation reads

\[
(1 - \lambda)(e^{\beta J-\beta \varepsilon} - \lambda) - e^{-\beta \varepsilon} = \lambda^2 - \lambda(1 + e^{\beta(J-\varepsilon)}) + e^{\beta(J-\varepsilon)} - e^{-\beta \varepsilon} = 0.
\]

Hence,

\[
\lambda = \frac{1}{2} \left(1 + e^{\beta(J-\varepsilon)} \pm \sqrt{(1 - e^{\beta(J-\varepsilon)})^2 + 4e^{-\beta \varepsilon}}\right).
\]

394
+ gives the Perron-Frobenius eigenvalue, so we can read off the free energy.

**Q27-4.** There is a long 1D lattice of adsorption points. Each adsorption point can accommodate at most one gas particle. The state of the lattice may be described by the adsorption pattern \( \{ \sigma_i \} \), where \( \sigma_i = 1 \) if the \( i \)th adsorption point is occupied, and \( \sigma_i = 0 \), otherwise. The adsorbed particles interact with each other if they are adjacent; the system energy \( \mathcal{H} \) may be described as

\[
\mathcal{H} = -\varepsilon \sum_i \sigma_i \sigma_{i+1}.
\]  (27.38)

This adsorbing 1D lattice is put in a very large box containing a gas that may be used as a chemostat for the adsorbing particles. The chemical potential of the gas may be assumed to be \( \mu \), a constant. We wish to determine the average coverage \( \theta \) of the lattice by the gas particles.

1. Write down the grand canonical partition function \( \Xi_M \) for the 1D lattice system of length \( M \) (do not try to compute the sum). Notice that the total number of the particles on the lattice may be written as \( N = \sum_i \sigma_i \). The temperature is maintained at \( T \). You may use the standard abbreviations as \( \beta \).

2. The grand canonical partition function of the length \( M \) 1D lattice described above may be written as

\[
\Xi_M = (1, 1)T^M a
\]  (27.39)

with the aid of a transfer matrix \( T \), where \( a \) is a certain 2 dimensional vector.

3. Find the limit \( q \)

\[
q = \lim_{M \to \infty} \frac{1}{M} \log \Xi_M.
\]  (27.40)

4. From \( q \) obtain \( \theta \).

**Soln.**

1. 

\[
\Xi_M = \sum_{\sigma_1, \ldots, \sigma_M \in \{0, 1\}} \exp[\beta \epsilon (\sigma_M \sigma_{M-1} + \sigma_{M-1} \sigma_{M-2} + \cdots + \sigma_2 \sigma_1) + \beta \mu N]
\]  (27.41)

\[
= \sum_{\sigma_1, \ldots, \sigma_M \in \{0, 1\}} \exp[\beta \epsilon \sigma_M \sigma_{M-1} + \beta \mu \sigma_M + \cdots + \beta \epsilon \sigma_2 \sigma_1 + \beta \mu \sigma_2 + \beta \mu \sigma_1].
\]  (27.42)

2. We rewrite (27.42) as

\[
\Xi_M = \sum_{\sigma_1, \ldots, \sigma_M \in \{0, 1\}} e^{\beta \epsilon \sigma_M \sigma_{M-1} + \beta \mu \sigma_M} e^{\beta \epsilon \sigma_{M-1} \sigma_{M-2} + \beta \mu \sigma_{M-1}} \cdots e^{\beta \epsilon \sigma_2 \sigma_1 + \beta \mu \sigma_2} e^{\beta \mu \sigma_1}.
\]  (27.43)
$e^{\beta \sigma_M \sigma_{M-1} + \beta \mu \sigma_M}$ can take the following values:

\[
\begin{array}{c|cc}
1 & 1 & 0 \\
\hline
0 & e^{\beta \epsilon + \beta \mu} & e^{\beta \mu} \\
\end{array}
\] (27.44)

Therefore, (27.42) may be written as

\[
\Xi_M = (1, 1) T^{M-1} \begin{pmatrix} e^{\beta \mu} \\ 1 \end{pmatrix}
\] (27.45)

with

\[
T = \begin{pmatrix} e^{\beta \epsilon + \beta \mu} & e^{\beta \mu} \\ 1 & 1 \end{pmatrix}
\] (27.46)

Since $M \gg 1$, you need not distinguish $M$ and $M \pm 1$.

(3) The characteristic equation for $T$ is

\[
\begin{vmatrix} e^{\beta \epsilon + \beta \mu} - \lambda & e^{\beta \mu} \\ 1 & 1 - \lambda \end{vmatrix} = (e^{\beta \epsilon + \beta \mu} - \lambda)(1 - \lambda) - e^{\beta \mu} = \lambda^2 - (1 + e^{\beta \epsilon + \beta \mu}) \lambda + 4e^{\beta \epsilon + \beta \mu} - e^{\beta \mu} = 0.
\] (27.47)

Therefore, the Perron-Frobenius eigenvalue is

\[
\lambda = \frac{1 + e^{\beta \epsilon + \beta \mu} + \sqrt{(1 + e^{\beta \epsilon + \beta \mu})^2 - 4e^{\beta \epsilon + \beta \mu} + 4e^{\beta \mu}}}{2} = \frac{1 + e^{\beta \epsilon + \beta \mu} + \sqrt{(1 - e^{\beta \epsilon + \beta \mu})^2 + 4e^{\beta \mu}}}{2}.
\] (27.48)

If $\mu \to -\infty$, then $\lambda = 1$ as expected.

(4) To obtain the coverage, we need the expected value of $N$.

\[
N = \frac{\partial}{\partial \beta \mu} \log \Xi_M
\] (27.49)

\[
\theta = \lim_{M \to \infty} \frac{N}{M}
\] (27.50)

Therefore, we conclude$^{322}$

\[
\theta = \frac{\partial}{\partial \beta \mu} q = \frac{\partial}{\partial \beta \mu} \log \lambda
\] (27.51)

$^{322}$You might worry about exchanging the limit and the differentiation, but since $\beta \mu$ is usually very negative, the convergence is swift, and the procedure is quite legitimate.
That is
\[
\theta = \frac{e^{\beta \epsilon + \beta \mu} + (2e^{\beta \mu} + e^{2\beta (\epsilon + \mu)} - e^{\beta (\epsilon + \mu)})/\sqrt{(1 - e^{\beta \epsilon + \beta \mu})^2 + 4e^{2\beta \mu}}}{1 + e^{\beta \epsilon + \beta \mu} + \sqrt{(1 - e^{\beta \epsilon + \beta \mu})^2 + 4e^{2\beta \mu}}}.
\] (27.52)

This indeed goes to 1 if \( \mu \) or \( \epsilon \) is large enough.
28 Symmetry breaking

Summary
* Spontaneous symmetry breaking and two important consequences are discussed.

Key words
(spontaneous) symmetry breaking, rigidity, Nambu-Goldstone boson

What you should be able to do
* Be able to explain intuitively why rigidity and NG bosons emerge upon spontaneous symmetry breaking.

28.1 Ordering means lowering the system symmetry
So far we discussed phase transitions. It is very often the transition between ordered and not-so-ordered states. Ordering means the system has less symmetry: in a gas phase molecules can sit anywhere, so the system has the full 3D translational and rotational symmetry, but if a crystal is formed (ordered!), we know the molecules cannot sit everywhere they wish; they must make a crystal lattice, so translational and rotational symmetries are lost.

28.2 How to describe the symmetry
The symmetry of a system may be understood through symmetry operations (Fig. 28.1).

It is clear that more highly symmetric objects allow more symmetry operations that keep the objects intact (invariant). The totality $G$ of the symmetry operations that keep an object intact is called the symmetry group of the object.\(^\text{323}\)

\(^{323}\)If $a, b \in G$, and if we write operating $b$ first and then $a$ next as the product $ab$, then $ab \in G$, so we can have an algebraic structure on $G$ (as illustrated in http://demonstrations.wolfram.com/C3vGroupOperations/). We know (i) the identity $e \in G$ and (ii) the inverse operation $a^{-1}$ of any operation $a \in G$ is again in $G$. Furthermore, (iii) $(ab)c = a(bc)$. If $G$ satisfies these three conditions, $G$ is called a group. If a subset $H \subset G$ is again a group, it is called a subgroup of $G$. Lowering of the symmetry of a system corresponds to restricting the original symmetry group to its genuine subgroup.
28.3 Spontaneous breaking of symmetry

If an equilibrium 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).\footnote{Z_2 group.}

Crystallization mentioned above is another example. The Hamiltonian of the system is something like

$$H = \sum_i \frac{p_i^2}{2m} + \sum_{i<j} \phi(r_i - r_j),$$

(28.1)

where $\phi$ is usually a binary interaction potential. Thus, the Hamiltonian has a full translation symmetry: nothing happens even if translation $r_i \rightarrow r_i + a$ is applied to all the particles; $H$ is invariant.\footnote{You might ask how the boundary of the system is taken care of. We take a huge system (eventually the thermodynamic limit), or we may impose a periodic boundary condition.} However, we believe this system crystallizes, losing its translational symmetry, if the interaction potential is something like the Lenard-Jones potential.\footnote{As already noted previously, we have not been able to prove this within statistical mechanics.} Thus, crystallization is a typical spontaneous symmetry breaking. In contrast to the 2-Ising model, the symmetry group in this case is continuous. That is, any $a$ is allowed or any small angle rotation can keep the Hamiltonian
intact. Really interesting phenomena due to spontaneous symmetry breaking occur if the broken symmetry is continuous.

28.4 Symmetry breaking in Heisenberg magnet
Consider a Heisenberg magnet as an example (cf. Fig. 28.2). In this case the system Hamiltonian
\[ H = -J \sum_{\langle i,j \rangle} \mathbf{s}_i \cdot \mathbf{s}_j \] (28.2)
has a full 3D rotational symmetry of spins.\(^{327}\) The disordered phase (paramagnetic phase) has no magnetization \( m = 0 \), so indeed the system is fully rotationally symmetric. However, below \( T_c \), when ferromagnetic order emerges, then \( m \) is a definite non-zero vector. Thus, the system symmetry is no more 3D rotational but only the 2D rotation around the axis parallel to \( m \) (in the spin space). Thus, the symmetry is lowered, a typical example of spontaneous symmetry breaking\(^{328}\) (Fig. 28.2).

Figure 28.2: 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.)

28.5 Consequences of symmetry breaking: rigidity
Now, take a Heisenberg magnet below its \( T_c \) with magnetization \( m \) being in the \(+z\)-direction. Let us choose one spin in front of us and rotate it by 90° to point in the \( O_3 \)-symmetry, needless to say, it is a continuous symmetry. The spins live on a lattice, so there is no spatial rotational symmetry. Do not confuse the rotations in the spin space and in the actual space.

\(^{327}\) In this case, the macroscopic states with different \( m \) are understood as distinct phases just as gas and liquid phases in fluids. If \( m \) changes to \( m' \), this is a first order phase transition between two distinct ordered phases.
+x-direction, and hold it. What happens? The spins around the rotated spin do not like this, because the interaction is energetically unfavorable. Thermal fluctuation occasionally flip them and align them to the held spin. Needless to say, then, these reoriented spins will have uncomfortable relations with further outside spins, BUT this outside relation is ‘better’; the central spin never moves, so the discomfort is steady, but in the outer layers it is ameliorated by thermal fluctuations. Thus, the flipped x-oriented domain gradually widens, and eventually the macroscopic magnet changes its direction of magnetization. Since this state has the same energy as the original state ($H$, the system Hamiltonian, is symmetric!), the final state will last forever, even if you stop holding the central spin. If we do not pay attention to what was actually happening between the two equilibrium states, what happens is just the rotation of the magnetization. This property—the whole system following the modification of its part—is called (generalized) rigidity.\footnote{The change in a ‘small part’ is in this case kept by an external means (by us). 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.}

The rigidity the most familiar to us is the rigidity of a solid. If we push one end of a solid, the other end also moves accordingly. You cannot do this for fluids. Only after translational symmetry is spontaneously broken can we have this ordinary rigidity of solid. If one end is twisted, the other end follows as well. This is due to the breaking of the rotational symmetry by crystallization.

Rigidity also occurs in 2D Ising model: if you flip the central spin and hold it, eventually the magnetization would change its sign. Thus, whenever symmetry is spontaneously broken, rigidity emerges. However, symmetry breaking of continuous symmetries is much more dramatic, because any local small change (which is impossible for discrete symmetry cases) propagates to the other end.

\section*{28.6 Nambu-Goldstone bosons: a consequence of breaking of continuous symmetry}

If the spontaneously broken symmetry is continuous, and if the system interactions are short-ranged, we have another universal feature: the \textit{Nambu-Goldstone bosons} (NG bosons). 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.

All possible symmetry broken phases (see Fig. 28.2) have the same energy, because they can be transformed into each other with an element of the symmetry group of the system Hamiltonian. Consider a 3D rectangular parallelepiped, and assume that
the phase changes continuously along its one axis (x-axis) (Fig. 28.3)

Figure 28.3: The situation in which the spin directions (magnetizations) in the planes perpendicular to an axis (x-axis) change gradually.

Let us estimate the needed energy for such deformation of the magnetization per cross section perpendicular to the x-axis. The spin interaction energy is given by the scalar product of spins, so if the angle between the neighboring spins is a small angle $\theta$, the energy increases by $1 - \cos \theta \propto \theta^2$ relative to the perfect parallel case. Suppose the spin-spin angle changes by $d\theta$ if they are apart by $dx$ along the x-axis. The energy change per $dx$ is proportional to $d\theta^2$, so the total energy change due to this twisting of spins from one end 0 to the other end $L$ is given by

$$\int_0^L (d\theta)^2 dx \simeq \int_0^L (d\theta) dx = \int_0^L \frac{d\theta}{dx} dx.$$  \hspace{1cm} (28.3)

Here, the $L/dx$ in the leftmost expression is the number of slices. The formula implies that if the total twist angle is $\theta$, then the total energy cost for this deformation is proportional to $\theta^2/L$. That is, if we can deform the system continuously, longer wave deformation (fluctuation) requires less energy to realize. Thus, the Nambu-Goldstone bosons become possible.

In the case of the Heisenberg ferromagnet, precession of spins can propagate as a wave (spin waves) and its quantum is called magnons. In a crystal the vibration due to the mutual displacement of lattice cells can propagate as a wave and its quantum is our familiar (acoustic) phonons. They are the NG bosons due to crystallization.

### 28.7 NG bosons do not exist for long-range interaction systems

However, as can be guessed from the above explanation, if there is a long-range interaction, then the energy required by a long-wavelength excitation may not vanish. This indeed happens in plasmas. Suppose we displace + charges relative to − charges as in Fig. 28.4. The Coulomb interaction energy between the charge density fluctuations does not decrease with distance (remember the parallel plate capacitor). Therefore, the excitation energy has a lower cut off and the long-wave frequency does
not converge to zero.\textsuperscript{330}

Figure 28.4: However far away the + and − charges are apart, in this case (recall a parallel plate capacitor) the Coulomb interaction between the separated charges does not decay.

28.8 Summary of symmetry breaking
We can summarize representative examples. Although not discussed superfluidity of $^4$He is also added.\textsuperscript{331}

<table>
<thead>
<tr>
<th>Broken Symmetry</th>
<th>solid</th>
<th>Heisenberg ferro</th>
<th>superfluid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>3D translational</td>
<td>rotational</td>
<td>phase</td>
</tr>
<tr>
<td>NG boson</td>
<td>3D periodicity</td>
<td>ferromagnetism</td>
<td>superfluidity</td>
</tr>
<tr>
<td>Rigidity</td>
<td>acoustic phonons</td>
<td>spin wave</td>
<td>second sound</td>
</tr>
<tr>
<td></td>
<td>rigidity</td>
<td>ferromagnetism</td>
<td>superfluidity</td>
</tr>
</tbody>
</table>

28.9 Symmetry breaking requires big systems
If the system is finite, there is no symmetry breaking.\textsuperscript{332} Fig. 28.2 implies the following difficulty: if we compute the partition function of a system as usual

$$Z = \sum e^{-\beta H},$$

(28.4)

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 is absolutely needed to make a rational and simple framework to understand spontaneous symmetry breaking.

\textsuperscript{330}Plasma oscillation is these excitations.

\textsuperscript{331}About this section, a strongly recommended reference is: P. W. Anderson, Basic Notions of Condensed Matter Physics (Westview Press 1984, 1997), Chapter 2.

\textsuperscript{332}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 given orientation is very long. However, mathematically, or theoretically, it is still not a true equilibrium state, so thermodynamic limit is taken.
28.10 What actually selects a particular symmetry broken state?
When the intrinsic symmetry is broken, how is a particular phase selected in the 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 (stray 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 performed after the thermodynamic limit; if performed 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.
29 First order phase transition

Summary
* How the first order phase transition becomes possible is explained.
* Even if phase transition occurs, the ensemble equivalence of statistical mechanics holds. That is, to study thermodynamics even with singularities, we can use any convenient statistical ensemble we like.

Key words
metastable state, unstable state, nucleation, spinodal decomposition

What you should be able to do
* Get familiar with the use of bifurcation diagram to understand phase transitions
Clearly recognize that $E$ is once continuously differentiable with respect to $S$, $V$ and other work coordinates.
* You must be able to illustrate why ensemble equivalence is all right.

29.1 First order phase transition example: nematic-isotropic liquid crystal transition
As already discussed briefly, in the case of liquid crystals, the ordering and volume change are coupled, and isotropic liquid-nematic liquid phase transition is (weakly) first order. As has already been mentioned, liquid crystal consists of slender molecules, which orient in the random directions at higher temperatures but tend to align at lower temperatures. This ordered phase is called the nematic liquid crystal phase. If we increase its temperature, due to thermal expansion, the distance between molecules increase slightly. This enhances disorganization of the molecular orientation, weakening molecular interactions further. This in turn enhances volume expansion and enhances disorder. In this way catastrophically order is lost, and a first order phase transition — nematic-isotropic phase transition — happens.

In this case, if there were no volume change, then the order would last more stably and the first order phase transition would be much closer to the second order phase transition or would become a second order phase transition itself.
29.2 Caricature model of first order phase transition
The above observation suggests a caricature model of first order phase transition within the mean field approximation. For a Ising magnet model, suppose that if the (magnitude of the) magnetization per spin $m$ (i.e., the order parameter) becomes smaller, $J$ in (26.23) decreases as illustrated in Fig. 29.1.

![Figure 29.1: 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 decrease.](image)

Let us review the mean field approach for a square lattice. Our starting point is the following equation

$$\langle s_0 \rangle = \langle \tanh[\beta J(s_1 + s_2 + s_3 + s_4)] \rangle. \quad (29.1)$$

The naivest mean-field approach is

$$m = \tanh 4\beta Jm \quad (29.2)$$

or

$$4\beta Jm = 4\beta J \tanh 4\beta Jm, \quad (29.3)$$

that is, we must solve

$$x = 4\beta J \tanh x. \quad (29.4)$$

We replace $J$ with the $J(x)$ in Fig. 29.1:

$$x = 4\beta J(x) \tanh x. \quad (29.5)$$

This modification is illustrated in Fig. 29.2.
29.3 Bifurcations exhibited by the caricature model
Let us study what happens if we lower the temperature. As $\beta$ increases, the curve in Fig. 29.2 becomes steeper and eventually crosses the diagonal line at three, and then five places as shown in Fig. 29.3; At $T_b$ new non-zero fixed points appear.

![Figure 29.3: 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.](image)

The stability of solutions may be read off from the bifurcation diagram Fig. 29.4. 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 Maxwell’s rule, which would choose a transition point (the thick vertical line) somewhere between $T_b$ and $T_X$ in Fig. 29.4.

29.4 Metastable and unstable states
In Fig. 29.4 the green curves denote stable solutions (roots) and red unstable solutions (in the sense that small perturbations added to the solution grow). Above $T_b$ without any question $m = 0$ disordered phase is the equilibrium phase, and below $T_X$ again without doubt the ordered (i.e., $m \neq 0$) phase is stable. Between $T_b$ and...
Figure 29.4: 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 thick vertical line in the figure; within the mean-field theory) we need a rule parallel to Maxwell’s rule.

$T_X$ the situation looks complicated. Thermodynamically, we expect there must be a phase transition from $m = 0$ branch to $m \neq 0$ branch somewhere between these two temperatures. The situation is analogous to the van der Waals gas; there should be a counterpart of Maxwell’s rule that determines the equilibriums phase transition point. Thus, if we come from the high temperature side slowly to $T_X$ a first order phase transition occurs at the vertical thick line position in Fig. 29.4.

If we rapidly cool the system, it is possible that we can stay on the green line below the phase transition point, which is the supercooled disordered phase, which is metastable: it is stable against small perturbations but it is not really globally stable (does not correspond to the global free energy minimum). If we heat the system in an ordered phase (say, $m > 0$ phase) gradually, at the phase transition point order is lost and the $m = 0$ phase appears. However, if we heat the system rapidly, we can continue to stay on the green curve, which is the superheated ordered state, and is metastable.

If we cool the disordered state really rapidly (temperature quench), then we could move the state on the green line left to $T_X$. This state is unstable, so it rapidly organizes into an ordered phase.

29.5 Phase ordering kinetics: nucleation and spinodal decomposition
How phases changes into each other is an interesting question both pure and materi-

---

333 Honestly speaking, we must make a mean-field approximation of the free energy and find the equilibrium condition.
als scientifically, because we could make various textures by arresting the transition process at an appropriate stage. When a metastable state orders, we expect seeds of ordered phases appear in the ocean of disordered phase as nuclei. The formation of nuclei is the rate-determining step. Once nuclei are formed, they grow rapidly and the phase transition is completed.

If a disordered phase is quenched into its unstable state, then immediately ordered domains appear everywhere in the space. However, ordered phases are usually not unique (in the figure we have \( \pm \) phases), so initially fine mosaic state is formed. Then, each domain of a particular ordered phase tries to increase its domain.\(^{334}\)

### 29.6 First order phase transition due to external field change

Phase transition can occur even if \( T \) is constant due to changes of other variables (say, \( P \) in the case of fluid; look at the \( PT \) diagram). Again, this phase transition can be understood intuitively with the aid of magnets.

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. This means one phase is no more a true equilibrium state but only a metastable state. Let us discuss the phase transition induced by this change with the aid of a mean field theory. Let us assume \( J \) is constant, and we consider (26.22), i.e.,

\[
m = \tanh(2d\beta J m + \beta h).
\]

(29.6)

Multiplying \( 2d\beta J \) and then adding \( \beta h \), we get

\[
2d\beta J m + \beta h = 2d\beta J \tanh(2d\beta J m + \beta h) + \beta h
\]

(29.7)

or we have only to study

\[
x - \beta h = 2d\beta J \tanh x
\]

(29.8)

To solve this equation we again use a graphic method (Fig. 29.5, the corresponding bifurcation diagram is in Fig. 29.6):

In Fig. 29.5 A \( \rightarrow \) E describes the effect of reducing magnetic field favoring the up phase while keeping the temperature \( T < T_c \). The corresponding bifurcation diagram (Left of Fig. 29.6) may be easier to understand. Below B once the down phase domain is formed, it is metastable (i.e., if it is large enough, it lasts for a very long time). The stability exchanges between the up and down phases at \( h = 0 \) can be

\(^{334}\)If the order parameter is conserved, it is called the spinodal decomposition.
Figure 29.5: Reducing \( h \) corresponds to \( A \rightarrow F \). If the magnetic filed intensity is positively large (A), the up phase is stable. Between A and B even if \( h \) is reduced virtually nothing happens. If \( h \) is reduced further a metastable down spin state (white disk on the negative domain) becomes possible. Also there is an unstable state (cross mark). If \( h \) is reduced, then the metastable down phase becomes stable, and the stable up phase becomes metastable. Look at the bifurcation diagram in Fig. 29.6 Left.

Figure 29.6: Bifurcation diagram for (29.8). Green curves denote locally stable solutions; thermodynamically (i.e., globally) stable or metastable states. Red portion denotes locally unstable solutions; thermodynamically unstable states.

understood intuitively. If \( h \) is further reduced (i.e., becomes larger in the opposite direction) at E the up phase becomes unstable. If there are remaining domains of the up phase, they disappear quickly.

Suppose the system is initially in the down phase. If an upward magnetic field (\( h > 0 \)) is applied, it becomes metastable, because the up phase is thermodynamically more stable (its free energy is less than that of the down phase). However, until B is realized, big enough down spin domains persist. If the magnetic field is suddenly increased to A, the down phase becomes unstable and goes into the up phase locally in avalanches.

The picture just explained applies to many first order phase transitions when the intensive variable is changed that is conjugate to a density that jumps at the first order phase transition. As can be guessed from the illustration in Fig. 25.7 for a fluid system (or a binary mixture system), pressure (or chemical potential) may be regarded as the intensive variable to induce first order phase transitions.
We know there is no phase transition above $T_c$ for fluids. This corresponds to the bifurcation diagram on the right of Fig. 29.6. Smoothly, just as the up phase turns into the down phase and vice versa, in the case of the fluid, very high density states may be converted into very low density states through changing the pressure.

### 29.7 Phase transition and ensemble equivalence

To conclude this introductory course, let us review the meaning of ‘ensemble equivalence’: you may use any convenient ensemble that can produce a certain thermodynamic potential (generalized free energy) to compute any thermodynamic potential (especially $E$ and $S$) you wish.

In these lectures it has been stressed that 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. Therefore, it is natural to guess that even if we compute the Helmholtz free energy $A$, we may not be able to obtain $S = S(E, X)$ in its generality. But, actually, it is not the case: from $A$ we can fully reproduce $E$. If $A$ is differentiable, of course we know the Gibbs-Helmholtz relation, but no differentiation is needed.\footnote{It is solely due to the convexity of $-A$ as we discussed briefly long ago.}

However, since the thermodynamic coordinates are privileged variables, we should lose something. Indeed, we lose some detailed information. Let us see what we can preserve and what we lose when we move from the thermodynamic coordinate system (in terms of $E = E(S, V, \cdots)$) to something else (in the illustration below, to $A = A(T, V, \cdots)$).

### 29.8 $E$ must always be continuously differentiable

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,

$$dE = TdS - PdV + xdX + \cdots,$$

so its continuous differentiability must be satisfied in the region of thermodynamic
Thus, internal energy must be a $C^1$ (= continuously differentiable) convex function of entropy and work coordinates. Consequently, the worst singularity is the loss of twice differentiability. For example, the constant volume specific heat can become not definable. We know at the critical point this indeed happens.

### 29.9 What if $E$ is not twice differentiable?

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.

**Figure 29.7:** When twice differentiability is lost: In (A) it is assumed that the second-order differentiability has a problem at a single point $a$. In (B) this happens at two points $a$ and $b$.

Fig. 29.7 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.

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. As we already know very well, 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 below $T_c$). In case (A) a crude sketch of the energy function cannot tell whether the transition is first order or higher.

---

$^{336}$ $P$ and $T$ never jump when we change thermodynamic variables. Why? It is a deep question.
29.10 What do we lose by Legendre transformation?

To understand the coexistence of two phases under constant temperature discussed 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 $A = \min_S [E - ST]$. If this is rewritten in the form standard to convex analysis, it reads $-A = \max_S [ST - E]$ (i.e., $E^* = -A$). Thus, the free energy is convex upward as a function of temperature (In Fig. 29.8 the convex function $-A$ is illustrated).

![Legendre transformation](image)

Figure 29.8: Legendre transformation $E$ to $A$ (or $-A$).

Fig. 29.8 Left is just the same as Fig. 29.7 (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.

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 single point by the Legendre transformation (Fig. 29.8). 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. We lose the information about the relative amount of coexisting phases by the Legendre transformation. However, it should be noted that from the right graph in Fig. 29.8, we can completely reconstruct internal energy as a function of thermodynamic coordinates by the inverse Legendre transformation $E = \max_T [ST - (-A)]$ (i.e.,
\( E^{**} = E \). This is the implication of the ensemble equivalence.