I. SYLLABUS

The course is taught on the basis of lecture notes which can be obtained from my web site (http://www.niu.edu/∼veenendaal/563.htm). These notes contain the following topics:

I. LITERATURE

II. INTRODUCTION
- A. History up to Newton
- B. Daniel Bernoulli and the foundations of statistical mechanics
- C. History from Newton to Carnot
- D. Thermodynamics
- E. Atomism

III. SYSTEMS OF NONINTERACTING ARTICLES OR SPINS

- A. Two-level spin system
- B. History of statistical mechanics: Clausius and Maxwell
- C. Thermal equilibrium and entropy
- D. History of statistical mechanics: Boltzmann
- E. Boltzmann factor
- G. Free energy
- H. Ideal gas
- I. Pressure
- J. The ideal gas law
- K. Gibbs' paradox
- L. History of statistical mechanics: Gibbs
- M. Link to classical partition function

IV. Quantum mechanics

- A. Planck's quantum hypothesis
- B. Many-particle wavefunctions
- C. The chemical potential
- D. Quantum distribution functions

V. Systems of interacting particles or spins

- A. Van der Waals gas
- B. Critical points
- C. The Ising model: mean field
- D. Relations between liquid-gases transition and magnetism
- E. Ising model: exact solution in one Dimension

VI. Applications

- A. Superfluidity in Helium II

A. Literature

Thermodynamics

R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, London, 1982).
A. H. Carter, *Classical and Statistical Thermodynamics* (Prentice Hall, Upper Saddle River, NJ, 2001).
R. L. Pathria, *Statistical Mechanics* (Butterworth-Heinemann, Oxford, 1972). F. Reif, *Statistical and Thermal Physics* (McGraw-Hill, Boston, 1965).
C. Kittel and H. Kroemer, *Thermal Physics* (W. H. Freeman and Company, New York, 1980)
H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1971).

Quantum mechanics

see, e.g., my lecture notes for 560/1, http://www.niu.edu/∼veenendaal/p560.pdf.

Background

R. D. Purrington, *Physics in the nineteenth century* (Rutgers University Press, New Bruncswick, 1997).
Wikipedia, http://en.wikipedia.org/wiki/Main_page.

II. INTRODUCTION

A. ***History up to Newton***

We start out this course by recapping in a very condensed form the history of thermodynamics and statistical mechanics. Throughout the centuries, the basic problems in natural philosophy (as science was called before it was split up into physics, chemistry, biology, etc.) can be brought back to several opposites: particles versus waves, discrete/quantized versus continuum. It was only in the late nineteenth and early twentieth century that these issues were resolved (at least, for the moment). Quantum mechanics shows that the particle and wave nature are both aspects of the same entity. Statistical mechanics demonstrates that the fact that nature is discrete is not at odds with the continuous character that we generally experience.

Many ancient philosophies included five elements to explain nature. The elements are earth, water, air (or wind), fire, and aether. The first three describe the three states of matter. In modern terms: solid, liquid, and gas. The study of these different phases is an essential part of thermodynamics. Fire is combustion, and essentially a chemical reaction, another important element in thermodynamics, although not really a part of this course. The fifth element, aether, is more complicated and we return to that in a while. In Greek philosophy these concepts were listed by Empedocles (ca. 450 BC), although the same concepts already appeared much earlier in Asia, where it forms the basis of Buddhism and Hinduism. The ideas of the five elements have had a profound influence on western civilization and philosophy, medicine, and science. In Greek philosophy, the ideas were worked out by Aristotle (384 BC-322 BC), one of the most important persons in western thought. Aristotle was a student of Plato and a teacher of Alexander the Great. Unfortunately, the effect on the progress of science was not always positive. We cannot blame Aristotle for all of this. First, he was a philosopher and not a natural philosopher, more interested in the "why" than the "how". Renaissance scientist, such as Galileo, suffered particularly from the fact that Aristotle's ideas were taken over by the catholic church, and turned into a dogma as opposed to a theory. Since we are indoctrinated by Newtonian thinking, it is difficult to consider ancient Greek thought without dismissing it directly. First, we have have to understand how revolutionary the ideas of Galileo, Newton, and others were. Nowadays, we make a big deal of Einstein imagining how it would be like to move at the speed of light, but the approach of is not unlike that in classical physics. The crucial breakthrough is to imagine the laws of nature in an ideal world. We find it completely natural that the Newton's laws are first discussed in an ideal world and that subsequently all the aspects of the real world (friction, drag, etc.) are introduced. For most Greek philosophers, this was unacceptable. Why first describe something in an unreal world and then try to describe the real world. For Aristotle, resistance was essential in his way of thinking. Newton's

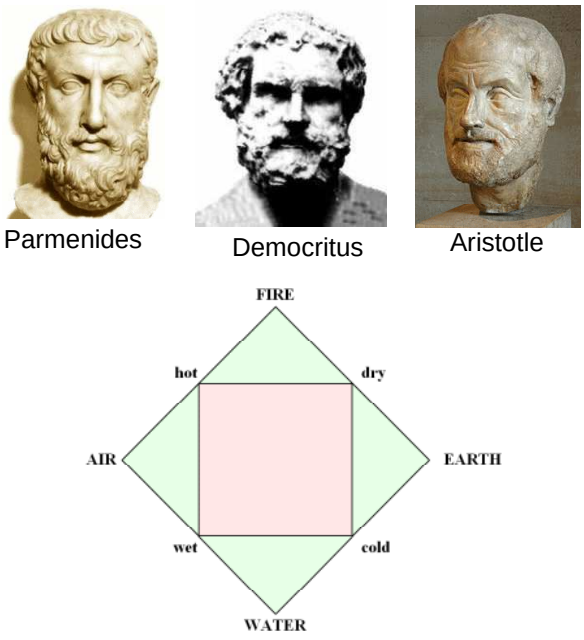

FIG. 1: The greeks Parmenides and Aristotle argues against the existence a void (or vacuum). The opposed the ideas of Leucippes and Democritus that everything was made out of indivisible particles (atoms) inside a void.

second law states, in one dimension, $F = ma = m\frac{d^2x}{dt^2}$. In the absence of forces, this gives $\frac{d^2x}{dt^2} = 0$ or $\frac{dv}{dt} = 0$ or $v =$ constant. This is Newton's first law. Newton's first law is contained in his second. Obviously, Newton was aware of this. However, he felt it necessary to state it separately (the question remains why, after more than four centuries, we still have to teach it that way). The reason is that it is unnatural and contradicting centuries of thought. It is unnatural since it is perfectly obvious to everyone that the velocity is not constant in the absence of a force. After a while, everything comes back to its natural state of rest. The Greeks had a more intuitive approach to laws that work in a real world. Objects have a nature. Why do rocks fall down? Because rocks are earth-like and it is in their nature to be with the rest of the earth. That's where they would like to move to. Feathers and leaves on the other hand are in between air and earth and fall much more slowly. Makes perfect sense. It also explains why rocks fall more slowly in water, because rocks are close to water than air. However, in the end everything comes back to rest. In fact, you require a force to move objects. Again, completely natural. To get somewhere you need to move your muscles. This whole concept of constant velocity in the absence of a force is completely ludicrous. In fact, some have credited Aristotle with the effective equation that the moving force can be written (in a very modern notation) as $F = Rv$, where R is the resistance and v is the velocity that you want (this is probably not what Aristotle meant, but often the interpretations of Aristotle's works are just as important as what was actually written). And probably a significant percentage of the world population would believe this law.

You might have heard this when discussing classical mechanics, but what does this have to do with thermodynamics. Well, things go horribly wrong when we go to Newton's ideal world or, as Aristotle would put it, the unreal world. Let imagine the motion in a void or vacuum. (And to give the Greeks credit, they did think about this. They noticed that their conception lead to contradictions and they were not afraid to confront these issues. They were open to discussion, unlike the centuries preceding Galileo). There are a number of problems. First of all, if it is in the nature of things to fall down, then what does an object do in a void. It would not like to go anywhere, because it would be in his nature to go to earth, which is absent in a void. So there is no motion in a void. Now what happens, when we apply a force. Well, since $v = F/R$, the velocity would be infinite and the displacement would be instantaneous. Since this is absurd, Aristotle against a vacuum and proclaimed his dictum of *horror vacui* or "nature abhors a vacuum" which was firmly states in his influential book *Physics*. Arguments against the void where already given more than a century before Aristotle, by Parmenides in ca. 485 BC. Parmenides argues against motion in a void, because a void is nothing and does not exist. If you move in there then something would come out of nothing, but *ex nihilo nihil fit* or nothing comes from nothing. Arguments against this were brought up by Leucippus, which written down and elaborated by Democritus (ca. 460 BC-370 BC). They proposed the, in our modern eyes, very appealing ideas that void was everywhere but filled with indivisible particles (*atomos*). The particles could have different shapes, size, and weights. Everything was made up of these tiny particles, or, in Democritus' words: "By convention sweet, by convention bitter, by convention hot, by convention cold, by convention color, but in reality atoms and void." Although revolutionary, the ideas were rejected by most Greeks. A nice idea to make up these particles, but nobody has ever seen them. Certainly ideas that still lived in the nineteenth century when Boltzmann and others were developing statistical mechanics. In addition, atomism, as it is commonly called, still requires moving in a void, which is impossible according to Aristotle.

This brings us to the fifth element. Surprisingly, many ancient cultures include this fifth element. Apparently, the four "earthy" ones, earth, water, air, and fire, were insufficient for a complete understanding of nature. Clearly, in Aristotle's world view the sky posed a problem. The moon and sun are obviously heavy entities and it would be in their nature to fall to earth. However, this does not happen, therefore they must be moving around in some other medium. This medium is called the aether. Obviously, they could not be moving around in a void, since that does not exist. The sky or the heavens is also where the God/Gods live and the aether is often also used to describe everything else that exists outside the material world, such as thoughts, ideas, love, etc. It is therefore quintessential to life itself (note that *quinta essentia* means fifth element). The word aether comes from "pure, fresh air" or "clear sky" in Homeric Greek and was the pure essence that the Gods breathed as opposed to the *aer* that regular mortals breathe. Aether had no qualities such as hot, cold, wet, or dry, as opposed to the other elements. In hinduism aether is called akasha. If was from the akasha that all the other four elements were made. Certainly, a very appealing concept in modern physics which claims that at the time of the Big Bang, the universe was created out of nothing. To understand the impact of these philosophies, if only suffices to realize that it was not until 1887, that modern science was finally able to start leaving the aether behind with the Michelson-Morley experiment.

It is against this background that the scientific work started in the renaissance. Centuries of thought had gone into the current understanding of the world. Many thoughts were wrong, but not entirely unreasonable. Rethinking of thermodynamics started when Galileo Galilei (1564-1642) was asked a question about siphon in ca. 1643. Siphon is just the Latin and Greek word for tube. We can use a continuous tube to drain liquid from a reservoir through a higher point to another reservoir. The final end of the tube has to be lower than the liquid surface in the reservoir, see Fig. 1. The principle of syphoning was already known to the Egyptians and the Greeks and described by Hero of Alexandria (10-70 AD) in his treatise *Pneumatica*. The common understanding is that, since the liquid flows out at the end of the

tube, a vacuum is created in the tube. However, since nature abhors a vacuum, liquid is sucked into the tube. In the middle ages, the technique was used to drain mines (and in this day and age, it is still used to steal the gas from your tank). In 1630, Galileo was asked by Giovan Battista Baliani of Genoa why a syphon he had contructed to carry water over a 21 meter high hill failed. Galileo had ascertained that there were apparently limitations in the vacuum force and the maximum intermediate height was about 18 braccia (around 11 meters). Galileo's interpretation, although, as we know now, being incorrect, stimulated further research by Evangelista Torricelli (1608-1647) and others. The experiment was simple. Instead of using water, Torricelli used a much heavier liquid: mercury. He filled a tube with mercury, closed the top with his finger and put it upside down in a basin and removed his finger. The mercury dropped and stopped at a height of 76 cm. Torricelli arrived at two important and revolutionary conclusions. First, he had created a vacuum inside the tube. Second, he was convinced it was not the sucking force of the vacuum holding the mercury in the tube, but the weight of the air pushing the mercury in the tube. Additional experiments by, among other, Blaise Pascal (1623-62), Gilles Personne Roberval (1610-1675), and Robert Boyle (1627-1691) demonstrated the validity of his arguments. Performing the experiment in vacuum (the void within a void) showed that the absence of air resulted in the complete emptying of the tube. Coincidentally, Torricelli had also invented the barometer. This was followed by Blaise Pascal demonstration that the atmospheric pressure depends on altitude in the famous experiment in 1648 on the Puy-de-Dôme.

After Torricelli's demonstration of the vacuum, Otto van Guericke was the first to create a vacuum using a simple vacuum pump consisting of a piston and a cylinder with one-way flap valves in 1650. He used this to suck the air out of two copper hemispheres with mating rims which were then tightly held together by the air pressure of the surrounding atmosphere. This was famously demonstrated in front of emperor Ferdinand III and the Reichstag in Regensburg in 1654 where two teams of 15 horses were unable to separate the two hemispheres. This experiment was repeated in Magdeburg in 1657 and the experiment is also know as that of the Magdeburg hemispheres. The vacuum pump was further perfected by Robert Boyle and and his assistant Robert Hooke (1635-1703). Among other things, they demonstrated that there is no sound in vacuum, demonstrating that sound needs a medium to propagate. In addition, vacuum helped classical mechanics. Although Galileo had already postulated that object fall in the same way irrespective of their weight, shape, and size, this was only convincingly demonstrated by actually performing the experiments in vacuum.

Robert Boyle was an influential natural philosopher. He could be described as a gentleman scientist, being born in Ireland as the seventh son and fourteenth child of Richard Boyle, the "Great Earl of Cork". He was highly educated and learned Latin, Greek, and French as a child, was sent to Eton College. After three years in college, he embarked

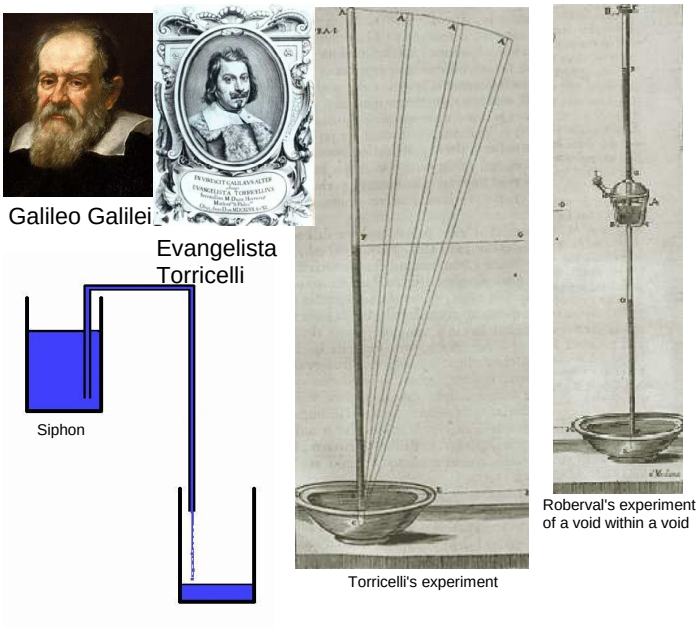

FIG. 2: Stimulated by Galileo Galilei's ideas that the reason that the intermediate height for a syphon was limited by the maximum strength of the vacuum force, Evangelista Torricelli performed his own experiments demonstrating the existence of a vacuum and the weight of air. Roberval's and others' reproduction of Torricelli's experiment in a vacuum (void within a void) demonstrated that it was indeed the weight of air keeping the mercury up since in vacuum the tube emptied.

on a *grand tour* with a French tutor (the term tourism is arrived from this. A typical rite of passage for wealthy British upper class men). He spent two years in Geneva and was in Florence in the winter of 1641. (Note, that Robert was only 14 at the time). It is worthwhile to remember that Galileo Galilei died in early 1642 within three miles of the great renaissance city. In 1645, his sick father left him Stalbridge manor in Dorset in Southern England. Boyle was part of the "invisible college", (first mentioned by Hooke in 1646) a group of scientists interested in pursuing ideas on this new natural philosophy. The group included John Wallis (1616-1703), who was an important figure in the history of calculus, crucial for the development of classical mechanics. He also introduced the symbol ∞ for infinity; Robert Hooke, Boyle's assistant, who played a role in many areas. Hooke is known for Hooke's law ($F = -Kx$), but also coined the term cell because his observation of plant cells reminded him of monks' cells; Sir. Christopher Wren (1632-1723). Wren is known to the general public as an architect. He was appointer the King's surveyor of Works in 1669, after the Great Fire of 1666 had destroyed two-thirds of London. Apart from designing St. Paul's cathedral, he was responsible for the rebuilding of 51 churches. He also contributed to astronomy and geometry. They corresponded through letters, before the existence of scientific journals. The informal and somewhat secretive societies formed the seeds for the foundation of the Royal Society in 1660. Robert Hooke served a Curator of Experiments, Wren was a co-founder and served as president. It had a great influence in the reporting and dissemination of science in the Philosophical Transactions of the Royal Society of London (which is the oldest English scientific journal and still exists. The oldest was French *Journal des Savants*). Robert Boyle was also influential through his extensive writings. His book The Sceptical Chymist (1661) is seen as a milestone in the field of chemistry. It denied the limitations to the classic four elements. It pushed chemistry to the same level as medicine and (yes!) alchemy. Boyle insisted that all theories must be verified experimentally.

Of more importance to us is that Boyle also wrote down one of the first quantitative laws in thermodynamics, known as Boyle's law. In modern form this is written as

$$PV = \text{constant}, \tag{1}$$

where P is the pressure and V is the volume. Although, this is certainly not a form that Boyle would have recognized. We are so used to our modern mathematical and physical tools that certain results look entirely trivial. Boyle wrote: *till further trial hath more clearly informed me, I shall not venture to determine, whether or no the intimated theory will hold universally and precisely, either in condensation of air, or rarefaction: all that I shall now urge being, that ... the trial already made sufficiently proves the main thing, for which I here allege it; since by it, it is evident, that as common air, when reduced to half its wonted extent, obtained near about twice as forcible a spring as it had before, so this thus comprest air being further thrust into half this narrow room, obtained thereby a spring about as strong again as that it last had, and consequently four times as strong as that of the common air.* It is clear that our modern mathematical tools greatly simplify our work and make it difficult for us to understand why it took so long for somebody to come up with something simple as PV =constant. Getting a nail in the wall looks easy when somebody had already invented the hammer, but try to imagine doing it without a hammer. Boyle did not invent this law, he wrote it down in 1662. It was actually discovered by two friends and amateur scientists Richard Towneley and Henry Power. The law was also discovered independently by Edme Mariotte (1620-1684) in 1676 (note that this is 14 years later!). In some countries, it is therefore known as the Boyle-Mariotte law. (Interestingly, as another example of the wide range of topics scientists studied in those days as opposed to the narrow-minded scientists of today, Mariotte was the first to describe the eye's blind spot. An interesting flaw in nature's/God's design of the eye. The optical nerves are attached to the receptors at the inside of the eye. This forces the optical nerves to exit the eye at some point where the are no optical receptor. Strangely enough, it is only cephalopods (squids and such) where nature/God got it right and let the optic nerve approach the receptors from behind, so that there is no need to break the retina). Combined with Charles' law (1787): $V/T =$ constant, where T is the temperature; Gay-Lussac's law (1802): $P \sim T$; Avogadro's principle (1811) $V/n =$ constant, where ν is the number of moles in the gas, this finally constituted the ideal gas law $PV = \nu RT$, where R is the gas constant, which was first stated by Clapeyron in 1834. That only took about 170 years.

However, we are drifting off and racing ahead. No overview can be complete without Isaac Newton (1643-1727). His phenomenal work *Philosophiae Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy, shortly known as the *Principia*) laid the basis for all classical mechanics to come. It also is a giant leap forward in the development of calculus which was firmly established by Newton and Gottfried Wilhelm Leibniz (1646-1716). Newton's achievements in calculus and classical mechanics are one of the most influential, if not the most influential, contributions in science. His contributions to classical mechanics should be well known to every student of physics. Newton also made significant contributions to optics. He advocated that light consisted of corpuscles, although he did not really stress this point. Although again this looks appealing from our modern photon concept of light, Newton's ideas are far removed from the quantum-mechanical framework. However, these ideas came under fire by scientists who believed in the wavelike nature of light such as Robert Hooke and the Dutch scientist Christiaan Huygens (1629-1965). This led Newton to withdraw from public life to about a decade. However, even though the wave representation

was a better approach to light at that time, Newton's influence was such that the particle theory of light dominated for the next century.

B. Daniel Bernoulli and the foundations of statistical mechanics

The Bernoullis were a Swiss family containing eight prominent mathematicians, including Jacob Bernoulli (1654-1705), who discovered the Bernoulli numbers (which are related to the Riemann zeta function and therefore to one of the famous unsolved problems in mathematics: the Riemann hypothesis); his brother Johann Bernoulli (1667-1748) solved the problem of the catenary or the hanging flexible cable. He was professor at the University of Groningen in the Netherlands, when his son Daniel Bernoulli (1700-1782) was born. His father pushed him to become a business man, but in the end, he choose mathematics. The relationship with his father significantly deteriorated when they both entered a scientific contest from the University of Paris. His father could not bare the shame of having to share the first prize with his son. As a return favor, he even tried to steal Daniel's book *Hydrodynamics* and recall it *Hydraulica*. Daniel Bernoulli also work on a new theory on the measurement of risks which was the base of the economic theory of risk aversion and premium and did significant work on statistics by analyzing the morbidity rates from small pox in 1766. Daniel Bernoulli also worked on flow and is the one behind the Bernoulli equation, which relates pressure, velocity and height in the steady motion of a fluid.

Daniel Bernoulli, in 1738, was the first to understand air pressure from a molecular point of view. He considered a vertical cylinder closed at the bottom and with a piston and a weight at the top. The piston's weight was supported by the air pressure. Or in Bernoulli's words: "*let the cavity contain very minute corpuscles, which are driven hither and thither with a very rapid motion; so that these corpuscles, when they strike against the piston and sustain it by their repeated impacts, form an elastic fluid which will expand of itself if the weight is removed or diminished...*". Brilliant! Unfortunately, very few accepted this view. Most people were convinced that gas molecules stayed more or less in the same place repelling each other held together by the aether. Let us derive a modern version of the kinetic theory. Consider a box in which particles are moving with a velocity v (this can be generalized to average velocities, but for simplicity let us assume that the velocity is constant). The momentum p of a particle moving towards the wall is then $p = mv$. If it bounces of the wall elastically, the momentum after the colission is $p = -mv_x$. Its momentum change is then $-2mv$, so the momentum transferred to the wall is $2mv_x$. If the density of particles is $n = N/V$ then the total number of particles hitting the wall in a time Δt is $\frac{1}{2}nAv_x\Delta t$, since half of the particles within a distance

Daniel Bernoulli

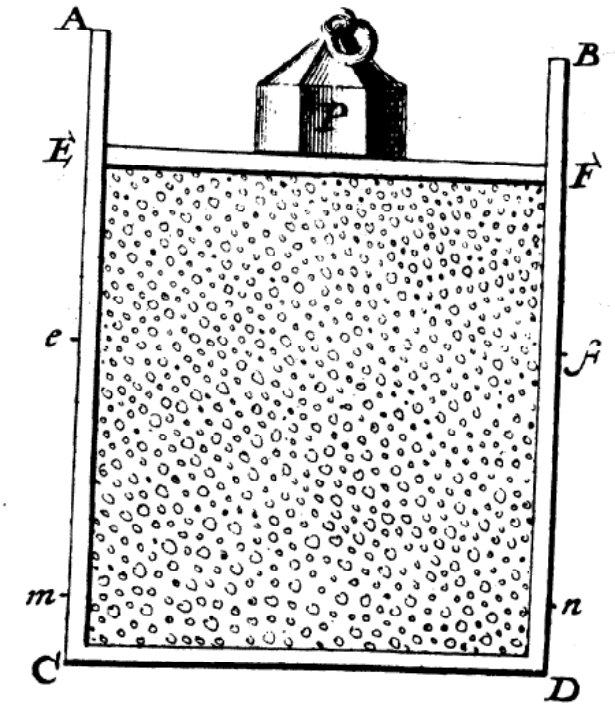

FIG. 3: Daniel Bernoulli and his illustration of the kinetic theory of gases from *Hydrodynamica* (1738).

$v\Delta t$ will hit the wall. The other half is moving in the opposite direction. The total momentum transferred to the wall is then $\Delta p = 2mv_x \times \frac{1}{2}nAv_x\Delta t = nAmv_x^2\Delta t$. The force exerted is then

$$F = \frac{\Delta p}{\Delta t} = nAmv_x^2. \tag{2}$$

The quantity v_x is one third of v^2, since $v^2 = v_x^2 + v_y^2 + v_z^2$. The pressure P is the force per surface area which is

$$P = \frac{F}{A} = nmv^2 = \frac{Nmv^2}{3V} \quad \Rightarrow \quad PV = \frac{1}{3}Nmv^2 = \frac{2}{3}N\overline{E}_{\rm kin}, \tag{3}$$

where $\overline{E}_{\rm kin}$ is the average kinetic energy per particle. This is Boyle's law if the term on the right hand side is constant. As said before this theory was neglected. This is not yet the ideal gas law. The relation between the kinetic energy and the temperature was first realized by John James Waterston (1811-1883) in the nineteenth century in an equally neglected book. Certainly, the title of the book, published at his own costs, *Thoughts on the Mental Functions* (1843) was not helpful. In modern notation, average kinetic energy can be written as $\overline{E}_{\rm kin} = \frac{3}{2}k_BT$, which we shall prove in this course. This yields the ideal gas law

$$PV = Nk_BT. \tag{4}$$

C. ***History from Newton to Carnot***

The eighteenth century saw the great developments of the first engine. The first engine was patented by Thomas Savery (1650-1715) in 1698. It was an early steam engine designed to pump water out of mines. Since the lifting effect of atmospheric pressure was limited to 30 feet, additional pressure coming from steam was used to increase the lift up to 50 feet. The machine used no pistons and was never applied successfully. Following Savery's design and further ideas by the Frenchman Denis Papin, Thomas Newcomen (1663-1729) made a more practical devise using a piston. Steam was injected into a cylinder from a boiler. The steam was then condensed and air pressure pushed down the piston, which was connected to the pump. Note that the real force is the atmospheric pressure since the pressure of the steam is limited. Further developments were made by James Watt (1736-1819) in the period 1763 to 1773. Watt's machine was the first to actually use the steam pressure above atmospheric. Watt noticed that Newcomen's engine

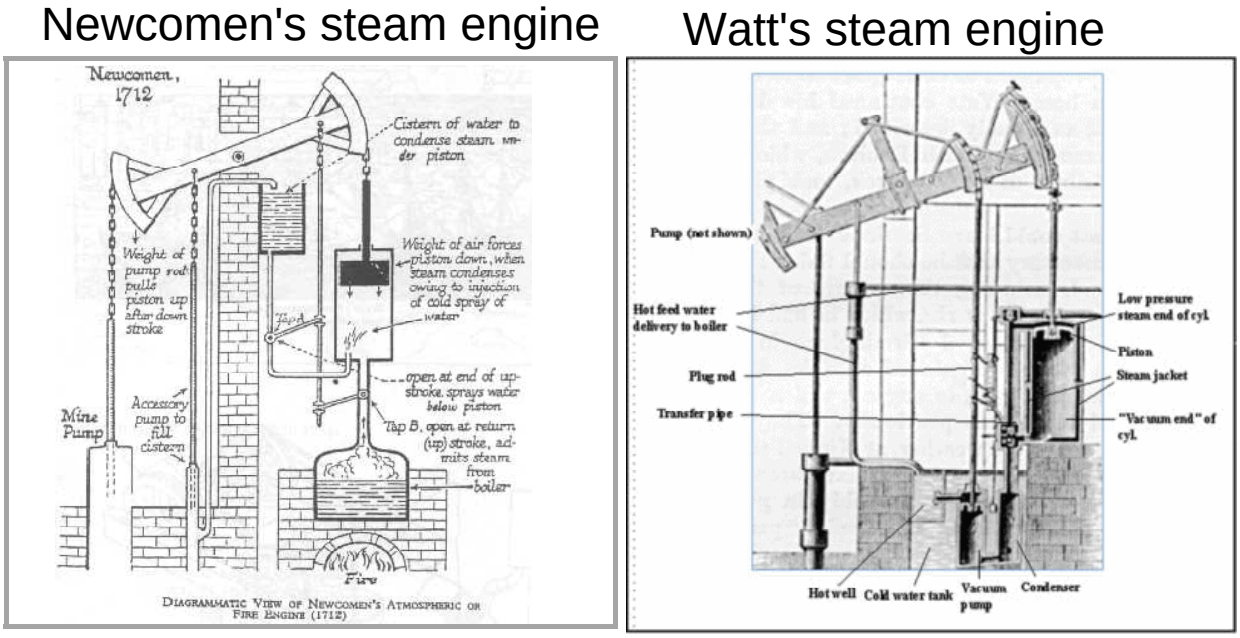

FIG. 4: The steam engines by Thomas Newcomen and James Watt to pump water out of mines. Note that the designs are very similar. The major difference is the introduction of the separate condenser. Other improvements in manufacturing also allowed Watt's machine to use steam pressure larger than atmospheric pressure.

was very inefficient. The major idea is the introduction of a separate condenser. When the piston reaches the top, the inlet valve is closed and the passage to the condenser is opened. This has the great advantage that the cylinder can remain hot and the condenser can remain cool, as opposed to having both processes take place in the same cylinder. This allows a great increase in speed. Another improvement was to have the downward motion of the piston done by low pressure steam as opposed to atmospheric pressure. Also the hot water in the condenser was removed by a vacuum pump and recycled into the boiler.

The theoretical understanding also made significant progress. However, the ideas of atoms and the corpuscule nature of light actually receded. One of the things we have to realize is that the concept of conservation of energy did not exist in the eighteenth century and the early nineteenth century. Early pioneers such as Leibniz already noticed in the later seventeenth century the conservation of the quantity (in modern notation) $\sum_i m_i v_i^2$. This he called the *vis viva* of the "living force". Other claimed that the quantity $\sum_i m_i v_i$ was the *vis viva*. Of course, we know that in elastic collisions both are conserved. However, for inelastic collisions, this is no longer correct, and some of the energy goes into heat. It was early on suggested that heat is related to the motion of atoms by natural philosophers, such as Newton, Boyle, and John Locke. Or in Boyle's words: "heat seems principally to consist in that mechanical property of matter we call motion". However, later on this was replaced by the caloric theory, which assumed that heat was fluid that flows from hotter to colder bodies or a weightless gas that moves in and out of the pores of solids and liquids. This confusion arose partly in the study of combustion. The seventeenth century had seen the development of the phlogiston theory by Johann Joachim Becher. The idea was that phlogiston was contained in combustible bodies and released during combustion. Nowadays we know that combustion is related to oxidation. However, the mistake is natural since many organic compounds get lighter during combustion since, e.g. water, CO, and CO_2, escape during the combustion. Since chemical reactions often release heat, the identification of heat as a substance appears more natural. This was corrected by the great French scientist/chemist Antoine Lavoisier (1742-1794) and laid the correct relation between combustion and oxidation. Materials therefore become heavier under combustion. Since at the same time their temperature rises. This appears in support of the caloric theory introduced in 1750 by the Scottish physicist Joseph Black (1728-1799). Black discovered latent heat. Latent heat is the heat that, for example, ice absorbs when it melts without change in temperature. Black noted that without latent heat a minute change in temperature would lead to the immediate melting of all the ice, in contrast with observation. Black concluded that heat was a liquid that combines with the material. The term latent comes from the Latin *latere*, to lie hidden. Lavoisier also introduced the concept of conservation of mass. Another statement that we all take for granted nowadays, but that was not at all obvious (we are leaving out radioactivity for the moment).

Another reason for the retreat of the atomic idea was the great success of continuous theories. Apart from the theory of Daniel Bernoulli, little else was achieved based on atomic principles. On the other had, significant progress was made in many areas using continuous models, especially through the development of differential equations. Joseph-Louis, comte de Lagrange (1736-1818) formulated, what is now known, as Lagrangian mechanics. Pierre-Simon Laplace (1749-1827) applied them in his theory of potentials. Joseph Fourier (1768-1830) applied an infinite series to study heat flow building on Newton's law of cooling. He was the first to treat heat conduction as a boundary value problem in the theory of partial differential equations. His idea that functions can be expanded in an infinite series had a great impact.

At the same time the foundations were laid for the theory of electricity and magnetism. Initial work was mainly electrostatics, such as the work by Charles-Augustin de Coulomb (1736-1806) in the 1780's on the strenght of the Coulomb interaction. Following Galvani's discovery of bioelectricity in 1783, Alessandro Volta invented the battery in 1800. The use of electrochemical reactions as opposed to the use of discharges from capacitors such as the Leyden jar (not to mention the outright dangerous methods of Benjamin Franklin) greatly facilitated research in this field. André-Marie Ampère (1775-1836) discovered the relationship between electricity and magnetism. This was followed by the very extensive work by Michael Faraday (1791-1867), discovering electromagnetic induction in 1831 and laying the basis for the electromotor. This finally culminated in the work of James Clerk Maxwell (1831-1879) combining electricity and magnetism in four equations (okay, he orginally had twenty), known today as Maxwell's equations. Electromagnetism was a continuous theory and certainly removed the ideas of corpuscular light proposed by Newton. Note that in the early nineteenth century electricity was also interpreted in terms of an electric fluid.

D. Thermodynamics

The early nineteenth also saw the foundations of thermodynamics. Obviously, few can claim to be the founder of anything, but the formulation of the Carnot cycle is generally considered an important milestone in the theory of thermodynamics. Nicolas Léonard Sadi Carnot (1796-1832) was a French physicist and military engineer. Note the dominance of French science in this period. One of the reasons for this is the French revolution and the institutionalization of science and the, at the same time, the foundation of the École normale and École polytechnique in the

1790's. This provided a great boost for science in France. Science is becoming a affair of the state, as is common nowadays. Carnot went to the École polytechnique. Many scientists were trying to understand the new steam engines that were turning the world upside down in an industrial revolution (note that, unlike in electromagnetism, science was more a follower of the technological developments). Carnot's work "Reflections on the motive power of heat and engines suitable for developing this power" (1824) is by many viewed as laying the foundation of modern thermodynamics. Carnot adhered to the caloric, although later in life he changed his view and became convinced that heat is related to the motion of atoms. In his work Carnot laid the basis for the second law of thermodynamics which states that the entropy of an isolated system not in equilibrium will tend to increase over time, appraoching a maximum value at equilibrium. A very basic consequence of that is that heat flows from a hotter to a cooler body. One does not observe that when two bodies are brought into contact that the cooler body becomes cooler and the hotter body hotter. Carnot was the first to consider full cycles of engines. He discovered that the work done by the engine is a function only of the the differential between the hold and cold reservoirs between which it operates. Later on, it was derived that the efficiency of an idealized engine is given by $(T_1 - T_2)/T_1$, where T_1 is the absolute temperature of the hotter reservoir.

Carnot's work was put on a firmer mathematical framework by Benoît Paul Émile Clapeyron (1799-1864). Clapeyron studied at the École polytechnique and École des Mines and left for Saint Petersburg to teach at the École des Travaux Public (note that the educated people in Russia spoke French and not Russian). He returned to Paris and supervised the construction of the first railway line connecting Paristo Versailles and Saint-Germain. Clapeyron devised a graphical representation of the Carnot cycle, developed the ideas of reversible processes, and more firmly established the Carnot principle, which is now known as the second law of thermodynamics.

The years following the death of its strong advocate Laplace in 1827, the caloric theory went into decline. Crucial early experiments (predating Laplace's death) were done by Benjamin Thompson (Count Rumford, 1753-1814) and Humphrey Davy (1778-1829). Thompson gave a famous paper to the Royal Society on the heat in cannon borings in 1798. Obviously, people had already realized that friction created heat. Thompson tried to measure the weight of the caloric fluid. The heat created in cannons is huge, however, the resulting loss in weight was zero. Davy presented related work on the the melting of ice due to friction in 1799. Although serious setbacks for the caloric theory, many scientists including William Thompson (Lord Kelvin) were not converted until the mid nineteenth century. Important steps were made by James Joule (1818-1889), an English physicist (and brewer) who studied the relationships between mechanical work and heat and laid the foundations for the law of conservation of energy. It is important to note that this important relationship (practically the basis of physics) was not established until the nineteenth century. One of the experiments performed by Joule is now known as Joule heating. In 1841, at the age of 23, Joule determined the heat loss I^2R by a current I flowing through a resistance R by placing the coil in water and carefully measuring the rise in temperature. He greatly generalized his work into a more general understanding of the interconvertability of different types of energy into each.

In 1850, the work of Carnot was put on a firmer basis by Rudolph Clausius. Clausius was born in the province of Pomerania in Prussia. In his Ph.D. work at the University of Halle he explained that the sky is blue due to the refraction of light. He then became professor of physics at the Royal Artillery and Engineering School in Berlin. In 1855 he moved to the ETH in Zürich. Later in life he also taught in Würzburg and Bonn. His most famous work is "On the motive power of heat, and on the laws which can be deduced from it for the theory of heat" in the Poggendorf Annalen in 1850. After several paper published by Clausius and William Thomson, this work presented the second law of thermodynamics for the first time in its complete form. Clausius defined entropy as $S = Q/T$, where T is the absolute temperature. He also introduced the name after the Greek word $\tau\rho o\pi\eta$ meaning transformation. He added the *en-* to make it sound similar to energy. Clausius also considered reversible and irreversible processes. Clausius also obtained an explicit form of the first law of thermodynamics in the familiar form $dQ = dE + pdV$.

Around the same time, William Thomson (1824-1907) devised the absolute energy scale, about which scientists had already speculated as early as Guillaume Amontons in 1702.

E. Atomism

Although we have seen that the ideas of atoms already existed in the support for the ideas of atoms was acually at a low point in the early and mid nineteenth century. Thermodynamics and many other theories worked very well in the absence of the assumption that everything is made up of small indivisible particles. In fact, continuum theories had made significant more progress than atomistic theories. The atomistic nature of matter was a largely unsupported hypothesis. In the early 1800s, apart from Bernoulli's work, atomism was more a philosophical speculation.

Atomism's resurgence started in chemistry. Chemistry had its foundations in the decades around 1800 with Priestley, Lavoisier, Dalton and Avogadro. It was John Dalton (1766-1844) who is generally credited with establishing chemistry on atomic principles. He was influenced strongly by Newton who assumed a static atomic model. A model of a gas

based on random motions was not proposed until 1738. Atoms were assumed to be massy and hard. The strong support for an atomistic model in chemistry is obvious. For example, two volumes of hydrogen combine with one of oxygen to produce water. Although some aspects were correct, other were less: Dalton assumed that matter was held together by attractive (gravitational) forces. The atoms were kept apart by the repulsive force of the caloric (the heat liquid). Other chemists were also crucial in propagating the atomic theory. The French chemists Pierre Louis Dulong (1785-1838) and Alexis Thérèse Petit (1791-1820), both at the école polytechnique, discovered that the specific heat of a body was directly related to its atomic weight, which is now known as the Dulong-Petit law.

In physics, the atomic principles were pushed by Davy and Faraday. They adhered to the point atomism ideas of written down by Roger Boscovich in his *Theory of Natural Philosophy* in 1763. Faraday was strongly influenced by his work in electrochemistry (together with John Daniell). He established the second law of electrochemistry in 1833 which asserts that "*the amounts of bodies which are equivalent to each other in their ordinary chemical action have equal quantities of electricity naturally associated with them.*" Or, the quantities of different elements deposited by a given amount of electricity are in the ratio of their chemical equivalent weights. Note that this implies a quantum of electricity, in effect the electron. Faraday had a problem with his picture of point-like atoms. If a solid is made of point-like atoms that do not touch each other, then what is causing the conduction. Certainly, not the space in between since how could there be insulators if the space is responsible? Obviously, Faraday was also concerned with the propagation of forces. Early on, he had believed that forces were transmitted by contiguous particles. Later on, he changed his ideas into the presence of force fields, while still clinging on to the presence of point-charge atoms in solids.

Although all the arguments in favor of the existence of atoms sound very convincing to our modern ears, many physicists remained unconvinced until the end of the nineteenth century. William Thomson wrote: "*The idea of an atom has been so constantly associated with incredible assumptions of infinite strength, absolute rigidity, mystical action at a distance, and indivisibility, that chemists and many other reseasonable naturalists of modern times, losing all patience with it, have dismissed it to the realms of metaphysics, and made it smaller than anything we can conceive.*" Despite this statement, Lord Kelvin went on to calculate the size of atoms and ended up with a surprisingly accurate values of 10^{-10} m, although he had a preference for vortex atoms. One might wonder how one can do electrostatics without the presence of point charges. However, here one can view the point charge as a mathematical devise to calculate continuous properties. There is no direct need to associate the point charge with an atom. It is against this background that the uphill battle for a discrete theory of mechanics starts employing the use of statistics.

III. SYSTEMS OF NONINTERACTING PARTICLES OR SPINS

A. Two-level spin system

Before discussing the historical developments of statistical mechanics, let us first develop some feeling for statistics when dealing with large numbers of particles. One of the simplest system that can be studied is a two-level system. This problem is equivalent to flipping coins.

See A. J. Kox, Eur. J. Phys. **18**, *139 (1997); Kittel-Kroemer chapter 1*

A more physically relevant two-level system is a set of spins in a magnetic field B. The splitting of energy levels under a magnetic field was first observed in the spectral lines of Sodium by the Dutch physicist Pieter Zeeman (1865-1943) in 1896 at the University of Leiden. His colleague at the university was the great Dutch physicist Hendrik Antoon Lorentz (1853-1928). Lorentz is known for the force named after him, which describes the force $\mathbf{F}$ acting on a particle with charge q moving at a velocity $\mathbf{v}$ in a magnetic field $\mathbf{B}$, $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$. Lorentz and Zeeman assumed that the particles were moving in an harmonic potential (comparable to a spring). This leads to the classical equations of motion

$$m\frac{d^2x}{dt^2} = -Kx + eB\frac{dy}{dt} \tag{5}$$

$$m\frac{d^2y}{dt^2} = -Ky - eB\frac{dx}{dt} \tag{6}$$

$$m\frac{d^2z}{dt^2} = -Kz \tag{7}$$

The solution of the latter is a simple harmonic motion $z = A\cos\omega_0 t$. Inserting gives a angular frequency $\omega_0 = \sqrt{K/m}$.

Two solutions are found for the xy plane

$$x = A\cos\omega_{\pm}t \quad \text{and} \quad y = \mp A\sin\omega_{\pm}t \tag{8}$$

Inserting gives

$$-\omega_{\pm}^2\cos\omega_{\pm}t = -\omega_0^2\cos\omega_{\pm}t \mp \frac{eB}{m}\cos\omega_{\pm}t \tag{9}$$

$$\pm\omega_{\pm}^2\sin\omega_{\pm}t = \pm\omega_0^2\sin\omega_{\pm}t + \frac{eB}{m}\sin\omega_{\pm}t \tag{10}$$

giving

$$\omega_{\pm}^2 \mp \frac{eB}{m}\omega_{\pm} = \omega_0^2. \tag{11}$$

Taking $\omega_{\pm} = \omega_0 \pm \Delta\omega$ gives

$$(\omega_0 \pm \Delta\omega)^2 \mp \frac{eB}{m}(\omega_0 \pm \Delta\omega) = \omega_0^2. \tag{12}$$

For small $\Delta\omega$, $(\omega_0 \pm \Delta\omega)^2 \cong \omega_0^2 \pm 2\omega_0\Delta\omega$. This gives

$$2\omega_0\Delta\omega = \frac{eB}{m}\omega_0 \quad \Rightarrow \quad \Delta\omega = \frac{eB}{2m} \tag{13}$$

When Zeeman and Lorentz evaluated the e/m ratio, they found a value much larger than expected. This is of course due to the fact that the mass of the electron is 9.1×10^{-31} kg whereas the mass of a ion is related to the mass of protons and neutrons 1.67×10^{-27} kg. Essentially they discovered the electron. Although they were puzzled about the value, they did not put a lot of emphasis on it, focusing on the interpretation of the spectral line shapes. However, this was 1896, a year before Joseph John Thomson published his famous experiments on the cathode ray tube. Fortunately Zeeman and Lorentz went on to win the Nobel prize for the Zeeman effect in 1902, whereas J. J. Thomson won it in 1906 for the discovery of the electron. For the energy, we can write

$$E = \hbar\omega - \mu_B B, \hbar\omega, \hbar\omega + \mu_B B, \tag{14}$$

where $\mu = \frac{e\hbar}{2m}$ is the Bohr magneton. Note that the identity $E = \hbar\omega$ was not known to Zeeman and Lorentz since this was proposed by Planck in 1900. In a quantum-mechanical notation we have

$$E = -\boldsymbol{\mu}\cdot\mathbf{B} \quad \text{with} \quad \boldsymbol{\mu} = \mu_B\mathbf{L}, \tag{15}$$

where $\boldsymbol{\mu}$ is magnetic moment and the $\mathbf{L}$ is the angular momentum. Along the z axis, L_z takes on the values $L_z = -1, 0, 1$ for a p electron ($l = 1$). We find three levels here, which is related to the fact that we are looking at the orbital part of the motion of the electron. We want to simplify it even further and consider the spin of the electron

$$E = -\boldsymbol{\mu}\cdot\mathbf{B} \quad \text{with} \quad \boldsymbol{\mu} = \mu_B g\mathbf{S}, \tag{16}$$

where $\mathbf{S}$ is the spin and $g \cong 2$ is the gyromagnetic ratio which is a relativistic effect. When choosing B along the z direction $\mathbf{B} = B\hat{\mathbf{z}}$, we obtain a two-level system

$$E = -\mu_B g B S_z = -\mu_B B\sigma, \tag{17}$$

where $S_z = \pm\frac{1}{2}$ and we take $\sigma = \pm1$. We take B positive, so that for the lowest energy level the spins are in the positive z direction.

For each spin, we have two possibilities: $\uparrow$ and $\downarrow$. We can write the Hamiltonian of this system as

$$H = -\mu_B B\sum_i \sigma_i \tag{18}$$

where for $\sigma_i = \pm1$, the index i runs over the spins (or the lattice sites, but since the spins are independent, their relative orientations is unimportant). In nature, one would like to have the minimum energy. This is given by $-N\mu_B B$, with N the number of spins. However, this is only true for a finite magnetic field at zero temperature. Let us take $B = 0$. In that case, we can still find the state with all the spins parallel, but there is only one such state. It

is already a lot more likely to find a state with one spin flipped of which there are N. How can we generalize this? We can write this as Newton's binomial

$$(\uparrow + \downarrow)^N = \sum_k \left(\begin{array}{c} N \\ k \end{array} \right) \uparrow^{N-k} \downarrow^k . \tag{19}$$

The coefficients given the probability g of finding a certain configuration

$$g_k = \left(\begin{array}{c} N \\ k \end{array} \right) = \frac{N!}{(N-k)!k!}. \tag{20}$$

This distribution is given in Fig. 5 for several values of N. Note that the factorials increase very rapidly. The probability of finding all spins parallel decreases quickly. For $N = 11$ it s 0.00098, for $N = 21$, is is already 9.5×10^{-7}. It may therefore seem like a wonder that with such a low probability we find all spins parallel. Yet we do, as we shall see in later sections. As we know from flipping coins the probability peaks around $\frac{N}{2}$. It is therefore convenient to express the spin of up and down spins as

$$N_\uparrow = N - k = \frac{1}{2}N + s; \qquad N_\downarrow = k = \frac{1}{2}N - s \tag{21}$$

For large N, this can be approximated using Stirling's approximation for the factorial

$$N! = \sqrt{2\pi N} N^N \exp[-N + \frac{1}{12N} + \cdots] \tag{22}$$

It is convenient to evaluate g by taking the logarithm

$$\ln \frac{N!}{N_\uparrow! N_\downarrow!} \cong \frac{1}{2}\ln 2\pi + (N + \frac{1}{2})\ln N - N - \left\{ \frac{1}{2}\ln 2\pi + (N_\uparrow + \frac{1}{2})\ln N_\uparrow - N_\uparrow + \frac{1}{2}\ln 2\pi + (N_\downarrow + \frac{1}{2})\ln N_\downarrow - N_\downarrow \right\} \tag{23}$$

$$= -\frac{1}{2}\ln 2\pi - (N_\uparrow + \frac{1}{2})\ln \frac{N_\uparrow}{N} - (N_\downarrow + \frac{1}{2})\ln \frac{N_\downarrow}{N} - \frac{1}{2}\ln N. \tag{24}$$

In the limit that $s \ll N$, we can write

$$\ln \frac{N_\uparrow}{N} = \ln \frac{\frac{1}{2}N + s}{N} = \ln \frac{1}{2}(1 + \frac{2s}{N}) \tag{25}$$

$$\cong -\ln 2 + \frac{2s}{N} - \frac{2s^2}{N^2}, \tag{26}$$

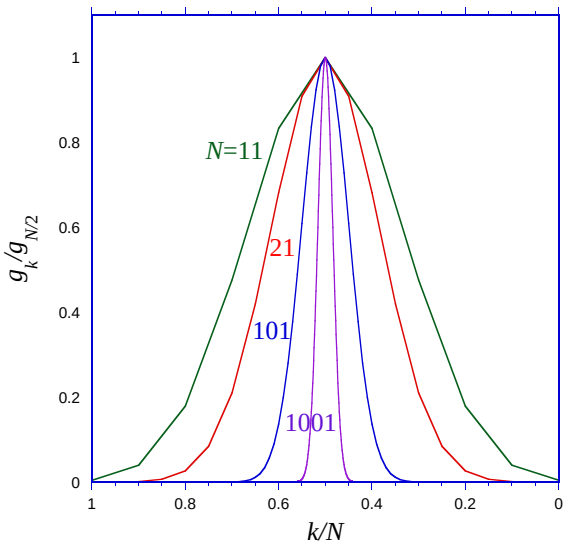

FIG. 5: The coefficients g_k in the binomial distribution as a function of k. Then x-axis is normalized to $N = 11, 21, 101, 1001$. The y-axis is normalized to the maximum value at $k = N/2$ with $g_{N/2} = 252, 184, 756, 1.0089 \times 10^2 9, 2.7028 \times 10^{299}$, respectively.

using the expansion $\ln x = x - \frac{1}{2}x^2 + \cdots$. Substituting gives

$$\begin{aligned}
\ln g &\cong -\frac{1}{2}\ln 2\pi N - (N_\uparrow + \frac{1}{2})(-\ln 2 + \frac{2s}{N} - \frac{2s^2}{N^2}) - (N_\downarrow + \frac{1}{2})(-\ln 2 - \frac{2s}{N} - \frac{2s^2}{N^2}) && (27)\\
&= -\frac{1}{2}\ln\frac{\pi N}{2} + N\ln 2 - (N_\uparrow - N_\downarrow)\frac{2s}{N} + (N_\uparrow + N_\downarrow + 1)\frac{2s^2}{N^2} && (28)\\
&\cong -\frac{1}{2}\ln\frac{\pi N}{2} + N\ln 2 - 2s\frac{2s}{N} + N\frac{2s^2}{N^2} && (29)\\
&= -\frac{1}{2}\ln\frac{\pi N}{2} + N\ln 2 - \frac{2s^2}{N} && (30)
\end{aligned}$$

or

$$g(s) \cong g(0)e^{-\frac{2s^2}{N}}, \tag{31}$$

with

$$g(0) = \sqrt{\frac{2}{\pi N}}2^N. \tag{32}$$

This distribution is known as a Gaussian distribution. Its integral from $-\infty$ to ∞ is 2^N, which is equal to the total number of possible arrangements of a system of N spins. The exact value of $g(0)$ is given by

$$g(0) = \frac{N!}{N_\uparrow N_\downarrow} = \frac{N!}{(\frac{1}{2}N)!(\frac{1}{2}N)!}. \tag{33}$$

However, the approximation is pretty good. For $N = 50$, the approximate value is 1.270×10^{14}, whereas the exact value is 1.264×10^{14}. In general, we are dealing with far larger numbers of spins. The value of the Gaussian is reduced to e^{-1} of its maximum value for

$$\frac{2s^2}{N} = 1 \quad \Rightarrow \quad s = \sqrt{\frac{N}{2}} \quad \text{or} \quad \frac{s}{N} = \frac{1}{\sqrt{2N}}. \tag{34}$$

For $N \cong 10^{22}$ (of the order of Avogadro's number) the fractional width is of the order of 10^{-11}. This is equivalent to the statement that when tossing coins you are much more likely to be close to 50% heads and 50% tails if you throw many times.

B. ***History of statistical mechanics: Clausius and Maxwell***

Following Carnot's paper in 1824 till the 1850s, thermodynamics developed rapidly. Thermodynamics is essentially a macroscopic theory and the conclusions are more or less independent from the underlying atomic nature. Early pioneers of the kinetic theory that tried to explain thermodynamic properties in terms of the motion of molecules were proposed by Daniel Bernoulli (as we saw earlier), John Herapath (1790-1868), and Joule. Much of this work was neglected (certainly, the fact that Herapath published his results in his own *Railway Magazine* did not help the dissemination of his research). Herapath found that the product PV was proportional to T^2 rather than T because he took the momentum proportional to the temperature and not the energy.

The modern theory of heat in terms of the motion of atoms and molecules was begun in the 1850s by August Krönig and the more profound work by Clausius. Clausius first paper on the subject was "The nature of the motion which we call heat" (1857) and was closely related to Krönig's work "The foundation of the theory of gases". Clausius next paper was entitled "One the mean lengths of the paths described by the separate molecules of gaseous bodies" (1858). The translation in the *Philosophical Magazine* in 1859 got the attention of James Clerk Maxwell (1831-1879). Clausius and Maxwell continued a fifteen year exchange in writing on the subject of the kinetic theory of gases. Their work initially focused on trying to understand the specific heat of gases. Carnot and Clapeyron had shown that the difference between the specific heats at constant pressure and volume is a constant, $C_P - C_V = \nu R$. Clausius explained the specific heat in terms of degrees of freedom of the atoms noting that we have three degrees of freedom for translation, but additional degrees of freedom for rotation and vibration of the molecules. Clausius also tried to describe the motion of atoms and molecules by the introduction of a mean free path between the collisions. To simplify the calculation he assumed that all the molecules move at the same velocity.

A huge leap forward in the development of statistical mechanics was made by Maxwell, who assumed that the speeds obeyed a probability distribution. Maxwell is one of the greatest physicists and is of course well known for his work in electricity and magnetism and aggregating all the equations into one coherent framework, now known as Maxwell's equations (note that he did not invent all of them, he only changed one of them). Although not crucial for his theory, Maxwell still believed that light propagation required a medium for the waves, known as the luminiferous aether (remember the fifth of the classical elements). Maxwell was born in Scotland and grew up on his father's estate in the Scottish countryside. Maxwell studied in Edinburgh and Cambridge and became professor at King's College London in 1860. In 1871, he became the first Cavendish professor in Cambridge. He oversaw the development of the Cavendish laboratory paid for by its founder the 7th Duke of Devonshire. Mawell oversaw the publication of the works of Henry Cavendish (1731-1810), son of the second Duke of Devonshire, most known for his determination of the density of the Earth (and the gravitational constant) and the composition of air and water. Maxwell died of abdominal cancer at the age of 48. Apart from electricity and magnetism, Maxwell worked on a wide variety of subjects, such as the composition of the rings of saturn where he showed that they consist of solid particles. He worked on color blindness. Recognizing that three colors are sufficient to produce all colors, he produced the first color photographs with the photographer Thomas Sutton.

Maxwell assumed that the volocities are distributed by a normal or Gaussian distribution, now known as the Maxwell (or Maxwell-Boltzmann) distribution. This approach was unacceptable to Clausius who, eventually, rejected the statistical approach. Maxwell went beyond gases and used his approach to derive coefficients for viscosity, diffusion, and conductivity.

C. Thermal equilibrium and entropy

see e.g. Pathria, p11, ; Kittel-Kroemer chapter 2.

Example: Thermal equilibrium between two spin systems.– Let us consider two two-level spin systems in thermal contact. The probability distribution is given by Eqn. (31). The probability of the combined system is

$$g_{12} = g(N_1, s_1)g(N_2, s_2) = g(N_1, 0)g(N_2, 0) \exp\left[-\frac{2s_1^2}{N_1} - \frac{2s_2^2}{N_2}\right]. \tag{35}$$

The total energy of the system is fixed. Since the total energy is related to the total spin by $E = -2\mu_B Bs$, this directly fixes the total spin to $s = s_1 + s_2$. Replacing $s_2 = s - s_1$ gives

$$g_{12} = g(N_1, s_1)g(N_2, s - s_1) = g(N_1, 0)g(N_20) \exp\left[-\frac{2s_1^2}{N_1} - \frac{2(s - s_1)^2}{N_2}\right]. \tag{36}$$

We would now like to find the value of s_1 that has the maximum probability. We can obtain this by taking the derivative of g_{12} with respect to s_1, or, more conveniently, the derivative of $\ln g_{12}$ with respect to s_1. The logarithm is given by

$$\ln g_{12} = \ln[g(N_1, s_1)g(N_2, s - s_1)] = \ln[g(N_1, 0)g(N_2, 0)] - \frac{2s_1^2}{N_1} - \frac{2(s - s_1)^2}{N_2}. \tag{37}$$

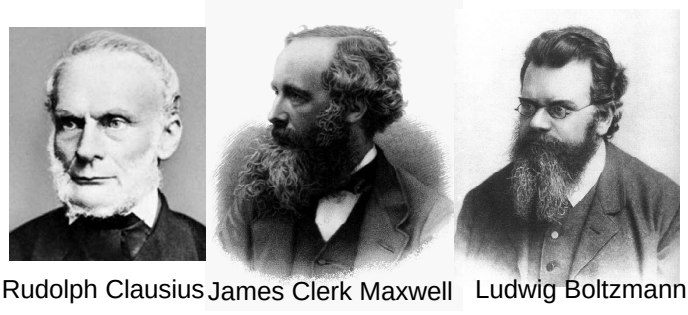

FIG. 6: The founders of the modern kinetic theory of gases and statistical mechanics: Rudolph Clausius, James Clerk Maxwell, and Ludwig Boltzmann.

The derivative is then

$$\frac{\partial \ln g_{12}}{\partial s_1} = -\frac{4s_1}{N_1} + \frac{4(s-s_1)}{N_2} \tag{38}$$

For the maximum the derivative is zero, which gives the condition for the maximum probability

$$\frac{4s_1}{N_1} = \frac{4(s-s_1)}{N_2}. \tag{39}$$

Solving for s_1 gives

$$s_{1,\max} = N_1 \frac{s}{N_1+N_2} = N_1 \tilde{s}, \tag{40}$$

where $\tilde{s}$ is the average s per site. This directly gives

$$s_{2,\max} = s - s_1 = N_2 \frac{s}{N_1+N_2} = N_2 \tilde{s}, \tag{41}$$

From the second derivative

$$\frac{\partial^2 \ln g_{12}}{\partial s_1^2} = -\frac{4}{N_1} - \frac{4}{N_2} < 0, \tag{42}$$

we see that this extremum is indeed a maximum. Basically, we see that if one were to remove the division between the two spin systems one would obtain one spin system with a homogeneous magnetization per site equal to $m = s/N$. Note that we can rewrite the product of the two Gaussian as a new Gaussian

$$g_{12} = g(N_1,0)g(N_2,0)\exp\left[-\frac{2(s_1-s_{1,\max})^2}{N_{1,\text{eff}}} - \frac{2s^2}{N}\right], \tag{43}$$

where

$$N_{1,\text{eff}} = \frac{N_1 N_2}{N_1+N_2} \to N_1 \quad \text{for} \quad N_1 \ll N_2. \tag{44}$$

We see that for $s_1 = s_{1,\max}$, the probability reduces to

$$g_{12} = g(N_1,0)g(N_2,0)\exp\left[-\frac{2s^2}{N}\right], \tag{45}$$

which is the probability of finding the system with a magnetization per site $m = s/N$. The dependence on s_1 describes the fluctuations of system 1. If system 2 is much larger than system 1 ($N_1 \ll N_2$), we can consider system 2 as a heat bath. We see that the fluctuation around $s_{1,\max}$ are Gaussian with a width proportional to $\sqrt{N_1}$.

1. *General treatment*

Let us consider systems 1 and 2 in equilibrium with each other. The systems are separate, so each has its own parameters such as N (for example, the number of particles or oscillators), V (Volume), and E energy. The systems are in thermal contact, so there is the possibility of exchange of energy, although we assume, for simplicity that N and V of 1 and 2 are fixed. The energies E_1 and E_2 are variable, but we have the restriction on the total energy

$$E = E_1 + E_2 = \text{const.} \tag{46}$$

The multiplicity of the combined system is

$$g(E) = \sum_{E_1} g_1(E_1) g_2(E-E_1). \tag{47}$$

The maximum is found by taking the derivative with respect to E_1,

$$\left(\frac{\partial g_1(E_1)}{\partial E_1}\right)_N g_2(E-E_1) + g_1(E_1)\left(\frac{\partial g_2(E-E_1)}{\partial E_1}\right)_N \times (-1) = 0 \tag{48}$$

or

$$\frac{1}{g_1}\left(\frac{\partial g_1}{\partial E_1}\right)_N = \frac{1}{g_2}\left(\frac{\partial g_2}{\partial E_2}\right)_N, \tag{49}$$

which we can also write as

$$\left(\frac{\partial \ln g_1}{\partial E_1}\right)_N = \left(\frac{\partial \ln g_2}{\partial E_2}\right)_N. \tag{50}$$

It therefore makes sense to define a quantity

$$\sigma = \ln g, \tag{51}$$

known as the entropy. The condition can now be rewritten as

$$\left(\frac{\partial \sigma_1}{\partial E_1}\right)_N = \left(\frac{\partial \sigma_2}{\partial E_2}\right)_N. \tag{52}$$

This condition is known in every day life as thermal equilibrium (or actually the inverse of it). The know that two systems are in equilibrium if the temperature are equal. The fundamental temperature is defined as

$$\frac{1}{\tau} \equiv \left(\frac{\partial \sigma}{\partial E}\right)_N. \tag{53}$$

The fundamental temperature is related to the normal temperature by

$$\tau = k_B T, \tag{54}$$

where k_B is a universal constant called the Boltzmann constant

$$k_B = 1.381 \times 10^{-23} \quad \text{J/K}. \tag{55}$$

Likewise the usual entropy is

$$S = k_B \sigma. \tag{56}$$

This gives Boltzmann's famous equation

$$S = k_B \ln g \tag{57}$$

We can rewrite the definition as

$$\frac{1}{T} \equiv \left(\frac{\partial S}{\partial E}\right)_N. \tag{58}$$

D. ***History of statistical mechanics: Boltzmann***

Boltzmann was born in Vienna the capital of the huge empire of Austria-Hungary. He studied at the university of Vienna. He received his Ph.D. in 1866 under the supervision of Joseph Stefan. It was Stefan who introduced Boltzmann to Maxwell's work. In 1869, at age 25, he was appointed Professor of Mathematical Physics at the University of Graz. He also visited Heidelberg working with Königsberger and Berlin where he met Kirchhoff and Helmholtz. In 1873, Boltzmann joined the University of Vienna. In 1876, he moved back to Graz, where he developed a lot of the theory of statistical mechanics. Among his students were Arrhenius and Nernst. Always restless and worried about his status, Boltzmann became chair of Theoretical Physics at the University of Munich in 1890. Austria, unhappy to loose on of its prominent scientists, weaned him back again to Vienna in 1893, where Boltzmann succeeded Stefan. However, Boltzmann had serious problems with his colleagues in particular Ernst Mach who was vehemently cricital of his work. One of Boltzmann's proof that was greatly criticized was his mechanical proof of the second law of thermodynamics, known as the H theorem. If shows that the entropy increases. However, the underlying physics laws are reversible in time. The paradox between microscopic reversibility and macroscopic irreversibility was pointed out by William Thomson in 1874 and Lohschmidt in 1876. This can still lead to vehement discussion since it basically determines

the arrow of time. It is true that it is possible to go to a state far from equilibrium (say all the molecules are in one corner of a space), but the probability of this happening is incredibly small. Other arguments against the H-theorem were brought up by Ernst Zermelo, a student of Max Planck, in 1896. Zermelo's used Poincaré's recurrence argument. Poincarré had demonstrated that any mechanical system is periodic, i.e. after a certain time it should return to its original state. This seems to be in contradiction with the second law of thermodynamics. This question is still not satisfactorily solved. One argument is that the recurrence time is so large that for all practical purposes it is infinite. Another argument focuses on the fact that the theorem is proven for isolated systems. However, on what time scale can a system be considered isolated. Random noise will destroy the recurrence. Therefore, we need to compare the recurrence time scale with the isolation time scale. The larger the system, the more difficult it is to keep it isolated and the more rapidly the recurrence time scale grows.

In 1900, Boltzmann moved to Leipzig only to return back to Vienna in 1902 after the retirement of Mach. In Vienna, Boltzmann not only taught physics but also philosophy. His lectures were extremely popular and overcrowded, despite being held in the largest lecture hall.

Boltzmann suffered from severe depression and attempted suicide several times. Nowadays, he would probably be diagnosed as bipolar. He reaction to criticism was unually strong. In 1906, during a summer holiday in Trieste he committed suicide in an attack of depression.

Although Boltzmann was a very capable experimentalists, his main achievements are in theoretical physics. Among Boltzmann achievements are the relationship between entropy and probability $S = k_B \ln W$, as Boltzmann wrote it, where W stands for *Wahrscheinlichleit* or probability. This equation is on Boltzmann's tombstone in Vienna. He also worked on dynamics and the equation describing the evolution of the distribution is named after him.

E. Boltzmann factor

Let us now consider a system in thermal equilibrium with a large reservoir. We want to know what the probability is of finding our sytem in a particular state with an energy E. Note that usually there is more than one state with a certain energy E. For example, for our spin system we want to know the probability of finding our system in one particular spin configuration. The total energy of the system and the reservoir is constant $E_0 = E_R + E$. Since we have specified our system, its probability is given by $g_s(E) = 1$. The total probability is therefore determined by the probability of finding the reservoir at an energy $E_R = E_{\rm tot} - E$,

$$g(E) = g_R(E_{\rm tot} - E) g_s(E) = g_R(E_{\rm tot} - E). \tag{59}$$

We can write that in terms of the entropy σ_R of the reservoir,

$$g(E) = \exp[\sigma_R(E_{\rm tot} - E)]. \tag{60}$$

Since the energy of the system is small compared to that of the reservoir, we can expand the entropy

$$\sigma_R(E_{\rm tot} - E) = \sigma_R(E_{\rm tot}) + \left(\frac{\partial \sigma_R(E_{\rm tot} - E)}{\partial E}\right)_{N,V} \times E + \cdots \tag{61}$$

$$= \sigma_R(E_{\rm tot}) - \left(\frac{\partial \sigma_R(E_{\rm tot} - E)}{\partial (E_{\rm tot} - E)}\right)_{N,V} \times E + \cdots \tag{62}$$

$$= \sigma_R(E_{\rm tot}) - \frac{E}{\tau} + \cdots \tag{63}$$

This gives a probability

$$g(E) \cong \exp[\sigma_R(E_{\rm tot}) - \frac{E}{\tau}] = g_0 \exp(-\frac{E}{\tau}). \tag{64}$$

Often we are only interested in relative probabilities of finding the system at, say, energies E_1 and E_2, for which we have

$$\frac{g(E_1)}{g(E_2)} = \frac{\exp(-\frac{E_1}{\tau})}{\exp(-\frac{E_2}{\tau})}. \tag{65}$$

The factor

$$w(E) = \exp(-\frac{E}{\tau}) = \exp(-\frac{E}{k_B T}) = \exp(-\beta E), \tag{66}$$

with $\beta = \frac{1}{k_B T}$, is known as the Boltzmann factor.

Example: temperature for a two-level spin system.– The energy for a two-level spin system is given by

$$E = -2s\mu_B B. \tag{67}$$

The probability can then also be written as a function of energy

$$g = g_0 e^{-\frac{2s^2}{N}} = g_0 \exp\left[-\frac{E^2}{2N\mu_B^2 B^2}\right] \tag{68}$$

The entropy is then

$$S = k_B \ln g = k_B \ln g_0 - \frac{k_B E^2}{2N\mu_B^2 B^2} \tag{69}$$

Taking the derivative with respect to E gives the temperature of the spin bath

$$\frac{1}{T} = \frac{\partial S}{\partial E} = -\frac{k_B E}{N\mu_B^2 B^2} \quad \Rightarrow \quad k_B T = -\frac{N\mu_B^2 B^2}{E} \tag{70}$$

Let us first consider that the reservoir is in the ground state with an energy $E = -N\mu_B B$, giving a temperature $T = \mu_B B$. This looks somewhat confusing since one expects the temperature to be zero when the system is in the ground state. Let us consider the Boltzmann factor in this limit:

$$w(E) = \exp\left(-\frac{E}{k_B T}\right) = \exp\left(-\frac{E}{\mu_B B}\right) \tag{71}$$

Note that E is the temperature of the whole reservoir. Fluctutation in the energy are of the order of $E \sim \delta N \mu_B B$. Since the total number of particles is of the order of $N = 10^{23}$, fluctations are of the order of $\delta N \sim \sqrt{N}$. The fluctuations in energy are therefore much larger than the temperature scale $k_B T$. We can look at it from a slightly different angle by considering the energy per site $\varepsilon = E/N$. A small change in entropy is now given by

$$w(\varepsilon) = \exp\left(-\frac{\varepsilon}{k_B T_0}\right) \tag{72}$$

with $k_B T_0 = \mu_B B/N$. We see that in the thermodynamic limit ($N \to 0$), $T_0 \to 0$. The probability of having excited states at T_0 therefore goes to zero. Note that the temperature in Eqn. (70) goes to infinity when $E \to 0$. $E = 0$ corresponds to a completely disordered sytem with $N_\uparrow = N_\downarrow = N/2$.

It is interesting to note that a system can also have negative temperatures. In this case, when $E > 0$. From the Boltzmann factor, we see that negative temperatures imply a larger occupation of the higher states, compared to the lower states. Although negative temperatures are not part of our daily experience, it is not impossible. For example, one could suddenly change the direction of the magnetic field and have all the spins pointing in the wrong direction. Actually, in our theoretical example this is easily achieved since the spins have no way of changing their direction. However, in real life this is somewhat more complicated. Any paramagnetic system rapidly decays due to the presence of a weak, but sufficiently strong coupling between the spins and the lattice. However, negative temperatures have been obtained for nuclear spins, which couple much more weakly to their surroundings, by Purcell and Pound, Phys. Rev. **81**, 279 (1951).

Summarizing, temperature in essence defines the excess energy in the system. Without our systems, the definition of temperature makes little sense.

F. Partition function

It is useful to define the function

$$Z = \sum_n w(E_n) = \sum_n e^{-\beta E_n}, \tag{73}$$

where the summation goes over all the possible states of the system. Z is called the partition function. This allows us to define a normalized probability

$$p(E_n) = \frac{w(E_n)}{Z} = \frac{1}{Z} e^{-\beta E_n}, \tag{74}$$

for which we have $\sum_n p(E_n) = 1$. Compare this with flipping coins and forget about temperature. We have two states (head and tail) and $w_{\rm head} = w_{\rm tail} = 1$. The partition function is the $Z = w_{\rm head} + w_{\rm tail} = 2$ and we can define the more useful quantities $p_{\rm head} = p_{\rm tail} = \frac{1}{2}$ This is an important result since it allows us to calculate the average energy of the system

$$E = \langle E \rangle = \sum_n E_n p(E_n) \tag{75}$$

$$= \frac{1}{Z} \sum_n E_n \exp(-\beta E_n) \tag{76}$$

Note that we can rewrite this as

$$E = -\frac{1}{Z} \sum_n \frac{\partial}{\partial \beta} \exp(-\beta E_n) = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}. \tag{77}$$

Or if we want to rewrite this in terms of temperature

$$E = -\frac{\partial \ln Z}{\partial \tau} \frac{d\tau}{d\beta} = \tau^2 \frac{\partial \ln Z}{\partial \tau} \tag{78}$$

using the fact that $\frac{d\beta}{d\tau} = \frac{d}{d\tau}\frac{1}{\tau} = -\frac{1}{\tau^2}$. Or in T,

$$E = k_B T^2 \frac{\partial \ln Z}{\partial T} \tag{79}$$

Example: Two-level spin system.–
See Kittel-Kroemer chapter 1

Let us return to our two-level spin system. The partition function is found by summing over all possible configurations of our system for a fixed number of spins N:

$$Z_N = \sum_{\{\sigma_i\}} \exp\left\{ \beta \mu_b B \sum_{i=1}^{N} \sigma_i \right\} \tag{80}$$

where $\{\sigma_i\}$ indicates all the permutations of the spins. We can multiply out the exponent

$$Z_N = \sum_{\{\sigma_i\}} \prod_{i=1}^{N} \exp(\beta \mu_B B \sigma_i) = \prod_{i=1}^{N} \sum_{\sigma_i = \pm 1} \exp(\beta \mu_B B \sigma_i). \tag{81}$$

Note that the product over the two different spin directions give all the possible configurations, as we saw in Eqn. (19). We can now write the partition function in terms of a single-spin partition function

$$Z_N = (Z_1)^N \quad \text{with} \quad Z_1 = \sum_{\sigma_i = \pm 1} \exp(\beta \mu_B B \sigma_i) \tag{82}$$

Z_1 can also be written as

$$Z_1 = e^{\beta \mu_B B} + e^{\beta \mu_B B} = 2 \cosh(\beta \mu_B B). \tag{83}$$

G. Free energy

See Kittel-Kroemer chapter 3

In the previous section, we saw how to obtain the energy from the partition function. However, when dealing with systems at finite temperature, there is a competition between entropy and energy. This is clear from our two-level spin system. To lower the energy, we would like to have all the spins parallel. However, at a finite temperature (which

is like adding energy to the system), spin will flip randomly. It far more likely to find a situation where not all spin are parallel, since the most likely configuration is that with 50% of the spins up and 50% down. In order to reflect this competition, it is useful to introduce a quantity known as the free energy

$$F = E - TS. \tag{84}$$

This quantity contains this balance between lowering the energy and an increase in disorder for higher temperatures. To lower F, we can lower E, but we can also increase S, i.e. go towards the state that is most probable.

Let us consider the differential at constant temperature

$$dF = dE - TdS. \tag{85}$$

Note that often we are dealing with problems at zero temperature (many quantum mechanics books consider only $T = 0$). In that case we have $dF = dE$. When the system is in the ground state, the energy E is in a minimum and therefore its differential is zero. At finite temperature and fixed volume F takes over the role that E has at zero temperature. Note that the temperature is defined as

$$\frac{1}{T} \equiv \left(\frac{\partial S}{\partial E}\right)_V \quad \Rightarrow \quad dE = TdS, \tag{86}$$

at constant temperature and volume. Therefore $dF = dE - TdS = 0$ and the system is in a minimum in thermal equilibrium at a constant volume.

For our spin system, we are interested in considering the situation where N and V do not change (later we will consider changes in N and V as well). A change in the free energy can be expressed as

$$dF = dE - TdS - SdT. \tag{87}$$

Note that a constant N and V, we can write

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} \quad \Rightarrow \quad dE = TdS. \tag{88}$$

Inserting this into the expression for dF gives

$$dF = -SdT \quad \Rightarrow \quad S = -\left(\frac{\partial F}{\partial T}\right)_{N,V}. \tag{89}$$

Giving for the free energy

$$F = E + T\left(\frac{\partial F}{\partial T}\right)_{N,V}. \tag{90}$$

The two terms with the free energy, we can rewrite using

$$-T^2\frac{\partial}{\partial T}\left(\frac{F}{T}\right) = -T^2\left\{F\left(-\frac{1}{T^2}\right) + \frac{1}{T}\frac{\partial F}{\partial T}\right\} = F - T\frac{\partial F}{\partial T}. \tag{91}$$

Combining this with the previous expression gives

$$E = -T^2\frac{\partial}{\partial T}\left(\frac{F}{T}\right)_{N,V} = k_B T^2\frac{\partial \ln Z}{\partial T} \tag{92}$$

or

$$F = -k_B T \ln Z. \tag{93}$$

Furthermore, we have

$$S = -\left(\frac{\partial F}{\partial T}\right)_{N,V} = k_B\left(\frac{\partial (T \ln Z)}{\partial T}\right)_{N,V}. \tag{94}$$

Two-level spin system see e.g. Pathria 3.9.

As we saw earlier, the partition function for a two-level spin system is

$$Z_N = (Z_1)^N = [2\cosh\frac{\mu_B B}{k_B T}]^N \tag{95}$$

From the partition function, we obtain the free energy

$$F = -k_B T \ln Z = -k_B T N \ln\left[2\cosh\frac{\mu_B B}{k_B T}\right], \tag{96}$$

the total magnetic energy of the system

$$\begin{aligned} E &= k_B T^2 \frac{\partial \ln Z}{\partial T} = k_B T^2 N \frac{\partial \ln\cosh\frac{\mu_B B}{k_B T}}{\partial T} && (97)\\ &= k_B T^2 N \frac{1}{\cosh\frac{\mu_B B}{k_B T}} \times \sinh\frac{\mu_B B}{k_B T} \times \left(-\frac{\mu_B B}{k_B T^2}\right) && (98)\\ &= -N\mu_B B \tanh\frac{\mu_B B}{k_B T}, && (99) \end{aligned}$$

and the entropy, which can be derived from

$$S = k_B \left(\frac{\partial(T\ln Z)}{\partial T}\right)_{N,V}, \tag{100}$$

but is more easily obtained by using

$$S = -\frac{F}{T} + \frac{E}{T} = k_B N\left\{\ln\left[2\cosh\frac{\mu_B B}{k_B T}\right] - \frac{\mu_B B}{k_B T}\tanh\frac{\mu_B B}{k_B T}\right\}. \tag{101}$$

The magnetization at constant temperature is defined as

$$M = \frac{\prod_{i=1}^{N}\sum_{\sigma_i=\pm1}\mu_B\sigma_i\exp(\beta\mu_B B\sigma_i)}{\prod_{i=1}^{N}\sum_{\sigma_i=\pm1}\exp(\beta\mu_B B\sigma_i)} = \frac{1}{\beta}\frac{\partial}{\partial B}\ln\left[\prod_{i=1}^{N}\sum_{\sigma_i=\pm1}\exp(\beta\mu_B B\sigma_i)\right] = k_B T\frac{\partial \ln Z}{\partial B} \tag{102}$$

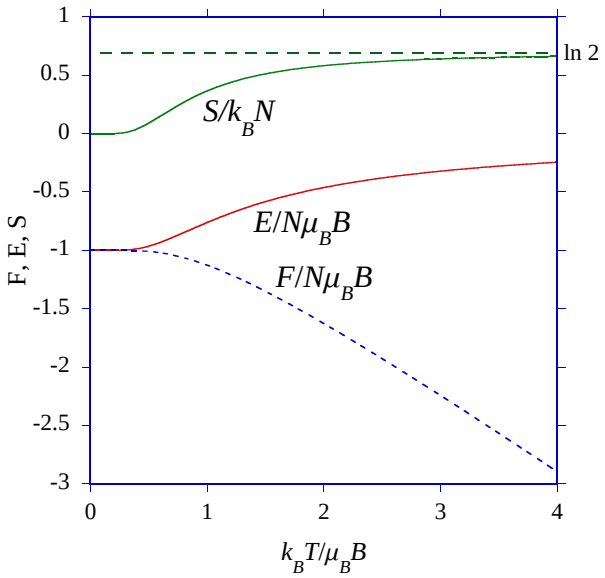

FIG. 7: The entropy S, magnetic energy E, and the free energy F. The latter two are given in $\mu_B B \equiv 1$.

or

$$M = -\left(\frac{\partial F}{\partial B}\right)_T \tag{103}$$

For the two-level system, this gives

$$M = k_B T N \frac{\partial}{\partial B} \ln\left[2\cosh\frac{\mu_B B}{k_B T}\right] = k_B T N \frac{1}{2\cosh\frac{\mu_B B}{k_B T}} 2\sinh\frac{\mu_B B}{k_B T} \times \frac{\mu_B}{k_B T} \tag{104}$$

$$= N\mu_B \tanh\frac{\mu_B B}{k_B T} \tag{105}$$

Note that the energy is directly related to the magnetization by $E = -MB$.

The behavior of the entropy S, the magnetic energy E, and the free energy F are given in Fig. 7.

Let us consider some limits. For $x \to 0$, we have

$$\sinh x = \frac{e^x - e^{-x}}{2} \to 0 \quad \text{and} \quad \cosh x = \frac{e^x + e^{-x}}{2} \to 1 \quad \text{and} \quad \tanh x = \frac{\sinh x}{\cosh x} \to 0 \tag{106}$$

For $x \to \infty$, we have

$$\sinh x = \frac{e^x - e^{-x}}{2} \to \frac{1}{2}e^x \quad \text{and} \quad \cosh x = \frac{e^x + e^{-x}}{2} \to \frac{1}{2}e^x \quad \text{and} \quad \tanh x = \frac{\sinh x}{\cosh x} \to 1 \tag{107}$$

The internal energy for $T \to 0$ is $E = -N\mu_B B$ (Note that here $x = \frac{\mu_B B}{k_B T}$. Therefore, in the limit $T \to 0$, $x \to \infty$). This corresponds to a completely aligned system, where every spin gain $-\mu_B B$ in energy. For $T \to \infty$ ($x \to 0$), $E \to 0$. In the absence of spin order due to the large temperature fluctuations, we have an equal amount of up and down spins and the average magnetic energy per spin approaches zero.

For the entropy, we find in the limit $T \to 0$ ($x \to \infty$)

$$S \cong k_B N \left\{\ln\left[2 \times \frac{1}{2}e^{\frac{\mu_B B}{k_B T}}\right] - \frac{\mu_B B}{k_B T}\right\} = k_B N \left\{\frac{\mu_B B}{k_B T} - \frac{\mu_B B}{k_B T}\right\} = 0. \tag{108}$$

This corresponds to a completely aligned state, which has a probability of 1 (or, actually, 1^N for the complete system) and therefore a zero entropy. In the limit $T \to \infty$ ($x \to 0$), the $\tanh x \to 0$ and the $\cosh x \to 1$ and we are left with $S = k_B N \ln 2$. The probability of the completely disordered state is $p = 2^N$, this gives $S = k_B \ln p = k_B N \ln 2$.

For the free energy we see that for $k_B T \ll \mu_B B$, the free energy is determined by the magnetic energy. At $T = 0$, we find

$$F = -k_B T N \ln 2 \times \frac{1}{2}e^{\frac{\mu_B}{k_B T}} - N\mu_B B = E(T = 0). \tag{109}$$

When the temperature becomes larger than the magnetic splitting, we will have a significant occupation of the spin state antiparallel to the magnetic field. The magnetic energy decreases and the entropy increases. For $k_B T \gg \mu_B B$, the free energy is determined by the entropy. For $T \to \infty$, we have

$$F = -k_B T \ln 2 = -TS(T \to \infty). \tag{110}$$

H. Ideal gas

See Kittel-Kroemer Chapter 3, Kestin-Dorfman Chapter 6.

Let us consider particles moving in a cubical box with lengths L. The potential is zero inside the box and infinite outside the box. We neglect spin and all kinds of other details of the particle. Inside the box, the wavefunctions satisfy the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) = \varepsilon\psi(\mathbf{r}). \tag{111}$$

The solution are well-known from every quantum mechanics course. We write $\psi(x,y,z) = X(x)Y(y)Z(z)$. We can now obtain a Schrödinger for each direction, such as

$$-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2} = \varepsilon_{n_x}X(x). \tag{112}$$

The solution is straightforward giving $X(x) = A_x \sin k_x x$. The wavenumbers have to satisfy the boundary conditions that the wavefunction is zero at the edge of the box. This gives $\sin(k_x L) = 0$ or $k_x = \frac{\pi}{L}n_x$ with $n_x = 1, 2, 3, \cdots$. We also have $\varepsilon_{n_x} = \frac{\hbar^2 k_x^2}{2m}$. For the total wavefunction we obtain

$$\psi(x,y,z) = A \sin\frac{n_x\pi x}{L}\sin\frac{n_y\pi y}{L}\sin\frac{n_y\pi y}{L}. \tag{113}$$

The energy is given by

$$\varepsilon_n = \frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2(n_x^2 + n_y^2 + n_z^2). \tag{114}$$

The partition function is now given by

$$Z_N = \sum_{\{n_i\}}\exp\left(-\frac{\sum_{i=1}^N \varepsilon_{n_i}}{k_BT}\right) = \sum_{\{n_i\}}\prod_{i=1}^N \exp\left(-\frac{\varepsilon_{n_i}}{k_BT}\right) = \prod_{i=1}^N\sum_{n_i}\exp\left(-\frac{\varepsilon_{n_i}}{k_BT}\right) = (Z_1)^N \tag{115}$$

where the summation n_i goes over all the possible permutations of N particles in all the quantum states. Again, as for the two-level system, we see that we can write the total partition function for the N particle system in terms of a one-particle system. Obviously, this is intimately related to the fact that the particles are independent.

$$Z_1 = \sum_{n_x,n_y,n_z}\exp\left[-\frac{\hbar^2\pi^2(n_x^2+n_y^2+n_z^2)}{2mL^2k_BT}\right] = \sum_{n_x,n_y,n_z}\exp[-\alpha^2(n_x^2+n_y^2+n_z^2)] = \prod_{i=x,y,z}\sum_{n_i}\exp(-\alpha^2 n_i^2) \tag{116}$$

wehre we simplified the expression using $\alpha^2 = \hbar^2\pi^2/2mL^2k_BT$. In the limit that the energy separation is small compared to k_BT (this should not be difficult since $\hbar^2\pi^2/2mL^2 = 3.8\times10^{-15}$ eV and room temperature is 25 meV), we can replace the summation by an integral, giving

$$Z_1 = \left(\int_0^\infty dn e^{-\alpha^2 n^2}\right)^3 = \left(\frac{1}{\alpha}\int_0^\infty dx e^{-x^2}\right)^3 = \frac{\pi^{3/2}}{8\alpha^3}, \tag{117}$$

using the substitution $x = \alpha n$. the fact that $\int_0^\infty dx e^{-x^2} = \frac{1}{2}\int_{-\infty}^\infty dx e^{-x^2} = \frac{1}{2}\sqrt{\pi}$. Using the definition for α

$$Z_1 = \frac{\pi^{3/2}}{8}\left(\frac{2mL^2k_BT}{\hbar^2\pi^2}\right)^{3/2} = \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}V = n_Q V = \frac{n_Q}{n}, \tag{118}$$

where we have used that the volume is $V = L^3$, which is related to the density by $n = 1/V$. The quantum concentration is defined by

$$n_Q \equiv \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}. \tag{119}$$

For Helium at atmospheric pressure at room temperature, $n \cong 2.5\times10^{19}$ cm^{-3}, and $n_Q \cong 0.8\times10^{25}$ cm^{-3}. Thus $n/n_Q \cong 3\times10^{-6}$. Whenever the ratio $n/n_Q \ll 1$, the gas can be considered to in the classical regime. An ideal gas is defined as a gas of noninteracting atoms in the classical regime.

We can calculate the energy from

$$E = k_BT^2\frac{\partial \ln Z}{\partial T} = k_BT^2N\frac{\partial \ln Z_1}{\partial T} = k_BT^2N\frac{\partial}{\partial T}\ln\frac{\pi^{3/2}}{8\alpha^2} \tag{120}$$

$$= k_BT^2N\frac{\partial}{\partial T}(\ln T^{3/2} + \text{terms independent of } T) = \frac{3}{2}Nk_BT \tag{121}$$

Of for the average energy per particle

$$\overline{E} = \frac{E}{N} = \frac{3}{2}k_B T. \tag{122}$$

This is the well-known result for the energy per atom of an ideal gas. Note that this implies that $\frac{1}{2}m\overline{v}^2 = \frac{3}{2}k_B T$ or $\overline{v} = \sqrt{3k_B T/m}$, where $\overline{v}$ is the average velocity. Note, that it is the average energy that is directly related to the temperature. At a certain temperature, the atoms in an ideal gas have a certain average energy irrespective of their mass. However, their average velocities do depend on the mass.

I. Pressure

In the previous section, we discussed the partition function and the energy of an ideal gas. Obviously, the first thing you remember from previous courses in thermodynamics about the ideal gas is the ideal gas law $PV = Nk_B T$, where P is the pressure. We therefore first need to introduce the concept of pressure.

Let us consider a system at a certain energy. We assume that the energy is a function of the volume of the system. We change the volume slowly by an external force in such a way that the systems remains in the same quantum state throughout the compression. For example, the system is still characterized by the same quantum states. Since the occupation of the states is essentially unchanged, the probability of finding this state is constant and therefore the change in volume occurs at constant entropy. The energy after the volume change is

$$\Delta E = E(V - \Delta V) - E(V) = -\frac{dE}{dV}\Delta V, \tag{123}$$

with $\Delta V > 0$. Let us consider for simplicity that our volume is a cube with the area of one face given by A. The volume change can then be written as

$$V = A(\Delta x + \Delta y + \Delta z). \tag{124}$$

The work done in the compression equals the change in energy

$$\Delta E = F(\Delta x + \Delta y + \Delta z) = \frac{F}{A}A(\Delta x + \Delta y + \Delta z) = P\Delta V, \tag{125}$$

where the pressure is defined as the force per surface $P = F/A$ and $\Delta x, \Delta y, \Delta z > 0$. Comparison with Eqn. (123) gives for the pressure

$$P = -\left(\frac{\partial E}{\partial V}\right)_S \tag{126}$$

Looking again at our system, we can seen that the number of states (and therefore the entropy) depends on the energy E and the volume V for a fixed number of particles. A change in entropy can therefore be written as

$$dS = \left(\frac{\partial S}{\partial E}\right)_V dE + \left(\frac{\partial S}{\partial V}\right)_E dV, \tag{127}$$

this gives the differential change in the entropy expressed in the differential changes dE and dV. Now we like to keep the entropy constant or $dS = 0$. Let us assume de can do that for a particular choice of $(dE)_S$ and $(dV)_S$. For this choice, we have

$$0 = \left(\frac{\partial S}{\partial E}\right)_V (dE)_S + \left(\frac{\partial S}{\partial V}\right)_E (dV)_S \quad \Rightarrow \quad 0 = \left(\frac{\partial S}{\partial E}\right)_V \frac{(dE)_S}{(dV)_S} + \left(\frac{\partial S}{\partial V}\right)_E \tag{128}$$

However, the ratio

$$\frac{(dE)_S}{(dV)_S} = \left(\frac{\partial E}{\partial V}\right)_S = -P \tag{129}$$

equal the partial derivative of E with respect to V at constant entropy. Using the definition of the temperature $\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_V$, we can rewrite the earlier result as

$$0 = \frac{1}{T}(-P) + \left(\frac{\partial S}{\partial V}\right)_E \tag{130}$$

or

$$P = T\left(\frac{\partial S}{\partial V}\right)_E. \tag{131}$$

Let us return back to the differential of the entropy. We can apparently rewrite this as

$$dS = \frac{1}{T}dE + \frac{P}{T}dV \tag{132}$$

or

$$TdS = dE + PdV \quad \text{or} \quad dE = TdS - PdV. \tag{133}$$

This important relation is known as the thermodynamic identity. We can identify this as the change in energy of the system dE is equal to the heat

$$dQ = TdS \tag{134}$$

added to the system and the work

$$dW = -PdV \tag{135}$$

done on the system.

Let us return to our definition of the free energy $F = E - TS$. The differential is given by

$$dF = dE - TdS - SdT \tag{136}$$

with the thermodynamic identity from Eqn. (133), this becomes

$$dF = -SdT - pdV. \tag{137}$$

We already knew that we could write the entropy in terms of the free energy

$$S = -\left(\frac{\partial F}{\partial T}\right)_V, \tag{138}$$

implicitly assuming all the time that the number of particles is fixed. We can also relate the pressure to the free energy

$$P = -\left(\frac{\partial F}{\partial V}\right)_T. \tag{139}$$

By using the definition of the free energy, we can also write

$$P = -\left(\frac{\partial E}{\partial V}\right)_T - T\left(\frac{\partial S}{\partial V}\right)_T. \tag{140}$$

The right hand side contains to terms which one might identify as energy pressure and entropy pressure. The energy pressure $-\left(\frac{\partial E}{\partial V}\right)_T$ is dominant for solids where interaction between the atoms cause a very large pressure making it very difficult to compress or expand solids. The entropy pressure $-T\left(\frac{\partial S}{\partial V}\right)_T$ is dominant in gases in gases and elastic polymers such as rubber. This might be somewhat counter intuitive, as one might expect that the change in energy is the most important contribution.

1. Maxwell relations

Using the expressions of the entropy and pressure in terms of partial derivatives of the free energy, we can derive thermodynamic relationship known as the Maxwell relations:

$$\left(\frac{\partial S}{\partial V}\right)_T = -\frac{\partial^2 F}{\partial V \partial T} = -\frac{\partial^2 F}{\partial T \partial V} = \left(\frac{\partial P}{\partial T}\right)_V. \tag{141}$$

That the partial derivatives of the entropy and the pressure are related to each other is far from obvious.

J. The ideal gas law

Let us return to the ideal gas. The total partition function can be written as

$$Z_N = (Z_1)^N = (n_Q V)^N, \tag{142}$$

Unfortunately, this is not the correct result. When dealing with the two-level spin system we could identify each spin with a certain lattice site. However, from quantum mechanics, we know that particles of the same kind are indistinguishable. We should therefore factor out all the possible permutations of the different particles among themselves. the correct form is therefore

$$Z_N = \frac{1}{N!}(Z_1)^N = \frac{1}{N!}(n_Q V)^N. \tag{143}$$

The free energy is given by

$$\begin{aligned} F &= -k_B T \ln Z_N = -k_B T N \ln(n_Q V) - k_B T \ln N! && (144)\\ &= -k_B T (N \ln V + N \ln n_Q - \ln N!). && (145) \end{aligned}$$

The pressure is now given by

$$P = -\left(\frac{\partial F}{\partial V}\right)_T = k_B T N \frac{1}{V}, \tag{146}$$

which directly leads to the ideal gas law

$$PV = N k_B T. \tag{147}$$

This is often written in terms of moles of gas, where one mole contains $N_A = 6.022 \times 10^{22}$ atoms or molecules.

$$PV = \nu R T, \tag{148}$$

where $\nu = N/N_A$ and $R = N_A k_B$.

We can also derive the entropy. It is convenient to approximate the factorial using Stirling's formula $\ln N! \cong N \ln N - N$

$$F = -k_B T (N \ln V + N \ln n_Q - N \ln N + N). \tag{149}$$

Before calculating the entropy, we should remember that the quantum density is a function of temperature $n_Q = a T^{3/2}$ The entropy is now

$$\begin{aligned} S &= -\left(\frac{\partial F}{\partial T}\right)_V = \frac{\partial}{\partial T}[k_B T (N \ln V - N \ln N + N) + k_B T N \ln a T^{3/2}] && (150)\\ &= k_B (N \ln V - N \ln N + N) + k_B N \ln a T^{3/2} + k_B T N \frac{3}{2}\frac{1}{aT} \times a && (151)\\ &= -k_B N \ln n + k_B N + k_B N \ln n_Q + \frac{3}{2} k_B N, && (152) \end{aligned}$$

which gives the Sackur-Tertrode equation

$$S = k_B N [\ln \frac{n_Q}{n} + \frac{5}{2}]. \tag{153}$$

K. Gibbs' paradox

In the previous section, we corrected the partition function for N particles by dividing by $N!$. This correction is directly related to Gibbs' paradox. Gibbs considered a system of weakly interacting particles (and noninteracting is about as weak as you can get). For such a system, we can divide the total system into two subsystems 1 and 2. Since

the particles are weakly interacting or noninteracting, we can also easily separate the energies of the two subsystems. We can then write the partition function as

$$Z = \sum_n e^{-\beta E_n} = \sum_{n_1,n_2} e^{-\beta(E_{n_1}+E_{n_2})} = \sum_{n_1} e^{-\beta E_{n_1}} \sum_{n_2} e^{-\beta E_{n_2}} = Z_1 Z_2 \tag{154}$$

Taking the logarithm gives

$$\ln Z = \ln Z_1 + \ln Z_1 \tag{155}$$

For the total entropy, we then have

$$S = k_B \ln Z + \frac{E}{T} = k_B \ln Z_1 + k_B \ln Z_2 + \frac{E_1}{T} + \frac{E_2}{T} = S_1 + S_2, \tag{156}$$

or the total entropy equals the entropy of the two subsystems.

Let us recalculate the entropy without the $1/N!$ correction factor

$$S = \frac{\partial}{\partial T}[k_B T N \ln V + k_B T N \ln a T^{3/2}] = N k_B (\ln V + \ln n_Q + \frac{3}{2}). \tag{157}$$

Suppose we subdivide our system into two equal systems S' and S'' with volumes $V' = V'' = V/2$ and $N' = N'' = N/2$. For a weakly interacting system, we would like to have $S = S' + S''$. However, let us calculate

$$S - S' - S'' = N k_B (\ln V + \ln n_Q + \frac{3}{2}) - 2 \times \frac{N}{2} k_B (\ln \frac{V}{2} + \ln n_Q + \frac{3}{2}) = N k_B \ln 2. \tag{158}$$

Obviously, this is not zero. However, when we assume the particles are indistinguishable, we have to divide our partition function by $N!$ to account for all the possible permutation among the particles. We then have

$$\ln \frac{Z}{N!} - 2 \times \ln \frac{Z'}{(N/2)!} \cong \ln Z - 2 \times \ln Z' - N \ln N + N - 2(-\frac{N}{2} \ln \frac{N}{2} + \frac{N}{2}). \tag{159}$$

The first term on the right hand side, we have just calculated, so we obtain

$$N \ln 2 - N \ln N + N + N(\ln N - \ln 2) - N = 0. \tag{160}$$

Introducing indistinguishability therefore removes Gibbs' paradox and we obtain $S = S' + S''$, as we would expect for a weakly interacting system.

1. Specific heat

See Reif Chapter 5.5-7

Specific heat defines how much the temperature T of a system changes when a certain amount of heat $dQ = TdS$ is added to the system. The specific heat at constant volume is defined as The specific heat at constant volume is defined as

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V = T \left(\frac{\partial S}{\partial T}\right)_V \tag{161}$$

Since dV is zero, the thermodynamics identity in Eqn (133) becomes $dE = TdS$ and we can write

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V \tag{162}$$

This can be evaluated straightforwardly using the expression for the average energy in Eqn. (122), which gives

$$C_V = \frac{3}{2} N k_B = \frac{3}{2} R \tag{163}$$

Later we want to express the specific heat in terms of the free energy. Let us start by using the expression of E in terms of the free energy from Eqn. (92),

$$E = -T^2 \left(\frac{\partial}{\partial T} \frac{F}{T} \right)_V . \tag{164}$$

We can therefore also write

$$C_V = -\frac{\partial}{\partial T} \left(T^2 \frac{\partial}{\partial T} \frac{F}{T} \right) = -\frac{\partial}{\partial T} T^2 \left(\frac{1}{T} \frac{\partial F}{\partial T} - \frac{F}{T^2} \right) = -\frac{\partial}{\partial T} \left(T \frac{\partial F}{\partial T} - F \right) \tag{165}$$

which gives

$$C_V = -T \left(\frac{\partial^2 F}{\partial T^2} \right)_V . \tag{166}$$

Alternatively, we can also determine the specific heat at constant pressure

$$C_P = \left(\frac{\partial Q}{\partial T} \right)_P = T \left(\frac{\partial S}{\partial T} \right)_P \tag{167}$$

Is there a relation between C_V and C_P? C_V gives the change in the system where we control the parameters T and V and we keep V constant (convenient for gases, less so for solids). For C_P, we control the parameters T and P and we keep the external pressure constant, for example, atmospheric pressure (most convenient for solids, doable for gases). Let us first consider the temperature and pressure as the independent variables and keep the volume constant, i.e. $S \equiv S(T, P)$. The amount of heat absorbed in an infinitesimal quasi-static process is then

$$dQ = TdS = T \left[\left(\frac{\partial S}{\partial T} \right)_P dT + \left(\frac{\partial S}{\partial P} \right)_T dP \right] . \tag{168}$$

Using the definition of C_P, one can rewrite this as

$$dQ = TdS = C_P dT + T \left(\frac{\partial S}{\partial P} \right)_T dP. \tag{169}$$

If the pressure is kept constant, $dP = 0$, we obtain the definition of C_P. However, in the calculation of C_V, T and V are used as independent variables. To express dQ in T and V, we express the differential dP in terms of dT and dV, giving

$$dQ = TdS = C_P dT + T \left(\frac{\partial S}{\partial P} \right)_T \left[\left(\frac{\partial P}{\partial T} \right)_V dT + \left(\frac{\partial P}{\partial V} \right)_T dV \right] . \tag{170}$$

The heat dQ absorbed when the volume is kept constant is obtained by setting $dV = 0$. Thus

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V = C_P + T \left(\frac{\partial S}{\partial P} \right)_T \left(\frac{\partial P}{\partial T} \right)_V . \tag{171}$$

This is a beautiful expression, but not very practical. For example, how are we supposed to determine experimentally $(\partial S/\partial P)_T$? To obtain more comprehensible quantities, we can make use of the Maxwell relations.

2. Gibbs free energy

We saw earlier that the free energy was given by $F = E - TS$. The differential is

$$dF = dE - TdS - SdT. \tag{172}$$

Using the thermodynamic identity $dE = TdS - pdV$, we obtain

$$dF = -pdV - SdT. \tag{173}$$

If we therefore have V and T as independent variables that we can keep constant we see the for $dT = 0$ and $dV = 0$, we have $dF = 0$ and therefore the free energy is in a minimum. However in the previous section, we were considering the situation where T and p were the indendent variables. Let us consider the Gibbs free energy defined as

$$G \equiv E - TS + PV. \tag{174}$$

The differential is

$$dG = dF + pdV - VdP = -pdV - SdT + pdV - VdP = -SdT - VdP. \tag{175}$$

We therefore find that the Gibbs free energy is at a minimum for constant temperature and pressure. Considering T and P as independent variables, i.e. $G = G(T, P)$, we can write

$$dG = \left(\frac{\partial G}{\partial T}\right)_P dT + \left(\frac{\partial G}{\partial P}\right)_T dP. \tag{176}$$

Since this has to be equal to Eqn. (175), we have

$$S = -\left(\frac{\partial G}{\partial T}\right)_P \tag{177}$$

$$V = \left(\frac{\partial G}{\partial P}\right)_T. \tag{178}$$

Since the cross derivatives are equal, we can obtain an additional Maxwell relation

$$-\left(\frac{\partial S}{\partial P}\right)_T = \frac{\partial^2 G}{\partial P \partial T} = \frac{\partial^2 G}{\partial T \partial P} = \left(\frac{\partial V}{\partial T}\right)_P \tag{179}$$

3. Specific heat: Continued

Earlier we were stuck with the expression

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_V = C_P + T\left(\frac{\partial S}{\partial P}\right)_T \left(\frac{\partial P}{\partial T}\right)_V, \tag{180}$$

which contained partial derivatives with little connection to experiment. From the Maxwell relation derived in the previous section, we can relate one derivative to a well-known quantity, namely the "volume coefficient of expansion", defined as

$$\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P. \tag{181}$$

This simply gives us the change in volume under the condition of constant pressure. Thus,

$$\left(\frac{\partial S}{\partial P}\right)_T = -V\alpha. \tag{182}$$

The derivative $(\partial P/\partial T)_V$ is also not readily determined since it requires that the volume is kept constant as a function of temperature. Although this is relatively easily done for gases the pressures produced by liquids and solids are very large. However, we can express the change in V in terms of P and T

$$dV = \left(\frac{\partial V}{\partial T}\right)_P dT + \left(\frac{\partial V}{\partial P}\right)_T dP. \tag{183}$$

At constant volume we have $dV = 0$. The quantity that we want to reexpress is given by

$$\left(\frac{\partial P}{\partial T}\right)_V = -\frac{\left(\frac{\partial V}{\partial T}\right)_P}{\left(\frac{\partial V}{\partial P}\right)_T}. \tag{184}$$

The numerator is again α. The denominator is another well-known quantity, since it is related to the change in volume due to a change in pressure at constant temperature. This is known as the isothermal compressibility, which is defined as

$$\kappa \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \tag{185}$$

$$\left(\frac{\partial P}{\partial T}\right)_V = -\frac{\alpha}{\kappa}. \tag{186}$$

Substituting the previous results into Eqn. (180), we obtain

$$C_V = C_P + T(-\alpha)\frac{\alpha}{\kappa} = C_P - VT\frac{\alpha^2}{\kappa} \tag{187}$$

or

$$C_P - C_V = VT\frac{\alpha^2}{\kappa}. \tag{188}$$

Example: Ideal gas
To evaluate the quantities for an ideal gas, let us start from the ideal gas law for ν moles of gas

$$PV = \nu RT. \tag{189}$$

We calculate first the expansion coefficient α. For constant p, we have

$$PdV = \nu R dT, \tag{190}$$

giving

$$\left(\frac{\partial V}{\partial T}\right)_P = \frac{\nu R}{P}, \tag{191}$$

hence

$$\alpha = \frac{1}{V}\frac{\nu R}{P} = \frac{\nu R}{\nu RT} = \frac{1}{T}. \tag{192}$$

Next, let us calculate the compressibility κ. For constant T, we have

$$PdV + VdP = 0, \tag{193}$$

giving

$$\left(\frac{\partial V}{\partial P}\right) = -\frac{V}{P}, \tag{194}$$

giving

$$\kappa = -\frac{1}{V}\left(-\frac{V}{P}\right) = \frac{1}{P}. \tag{195}$$

Thus we obtain for the difference

$$C_P - C_V = VT\frac{(1/T^2)}{1/p} = \frac{VP}{T} = \nu R. \tag{196}$$

Alternative way:
For an ideal gas we can write

$$dE = \left(\frac{\partial E}{\partial T}\right)_V dT. \tag{197}$$

For an ideal gas this is also true when the volume changes since the energy E is independent of the volume, so the partial derivative with respect to the volume is zero. Using

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V, \tag{198}$$

we can write

$$dE = C_V dT. \tag{199}$$

Inserting this in the thermodynamic identity gives

$$dQ = dE + pdV = C_V dT + pdV. \tag{200}$$

From the ideal gas law, we can find for constant pressure $pdV = \nu R dT$, leading to

$$dQ = C_V dT + \nu R dT \tag{201}$$

or

$$\left(\frac{\partial Q}{\partial T}\right)_P = C_V + \nu R. \tag{202}$$

The left hand side is the definition of C_P, giving as final result

$$C_P - C_V = \nu R. \tag{203}$$

L. ***History of statistical mechanics: Gibbs***

M. Link to the classical partition function

After doing the previous sections, you might say: "Hey, I don't like this at all. I don't picture an ideal gas as quantum mechanical particles in a box". Indeed, you might wander if this also works with classical noninteracting particles with kinetic energy $\frac{1}{2}mv^2$. We directly notice a problem if we want to write down the partition function. There are no states to sum over! We have continuous variables. Let us just move ahead and integrate over the relevant variables we can think of, namely $\mathbf{r}$ and $\mathbf{p}$. For one particle, we obtain

$$Z_1^c = \int d\mathbf{r} \int d\mathbf{p} e^{-\beta \frac{\mathbf{p}^2}{2m}}. \tag{204}$$

The integration over the volume is trivial; the integration over momentum, we can split up into components

$$Z_1^c = V \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z e^{-\beta \frac{p_x^2+p_y^2+p_z^2}{2m}} = V \left(\int_{-\infty}^{\infty} dp_i e^{-\beta \frac{p_i^2}{2m}}\right)^3 = V \left(\sqrt{\frac{2m}{\beta}} \int_{-\infty}^{\infty} d\xi e^{-\xi^2}\right)^3 \tag{205}$$

with $\xi = \sqrt{\beta/2m} p_i$. The integral over ξ is well known and gives $\sqrt{\pi}$. The total result is tehn

$$Z_1^c = (2\pi m k_B T)^{3/2} V. \tag{206}$$

Let us compare this with the classical partition function

$$Z_1 = \frac{1}{h^3} (2\pi m k_B T)^{3/2} V \tag{207}$$

where we used $\hbar = h/2\pi$. All right, close but not quite. There appears to be an h floating around. Then again, we would not quite expect Planck's constant to appear in a classical equation. The difference arises from the fact that in the classical partition function, we integrated over the variables, whereas in the quantum mechanical partition function, we summed over the states. Let us consider one particular quantum state. In momentum space, one quantum space takes up a volume of

$$dp_x dp_y dp_z = \hbar^3 dk_x dk_y dk_z = \hbar^3 \left(\frac{\pi}{L}\right)^3 = \frac{1}{8V} h^3. \tag{208}$$

However, there are eight equivalent octants corresponding to the equivalent quantum numbers $\pm n_x, \pm n_y, \pm n_z$. The total volume in real space is given by $L^3 = V$. Therefore, the total volume in phase space of one quantum state is

$$\Delta\mathcal{V} = 8 \times \frac{1}{8V} h^3 \times V = h^3. \tag{209}$$

To obtain the quantum mechanical partition function from the classical one, we need to divide by this volume to change from an integration over phase space to an effective summation over quantum states,

$$Z = \frac{1}{h^3} Z^c. \tag{210}$$

After doing this, we find a perfect agreement between the classical result for the ideal gas and the quantum-mechanical partition function.

IV. QUANTUM STATISTICS

A. Planck's quantum hypothesis

The quantum theory started with the blackbody radiation. Despite the fact that it is a rather complicated phenomenon, it is worthwhile to have a look at the origins of quantum mechanics. The aim is to find the energy dependence of the emitted radiation of a perfectly absorbing body (it is black). We can create this, for example, by making a box with a small hole. The experimental energy dependence is given in Fig. 2.1 of Liboff. In fact, a lot can already be said by thermodynamic arguments. Wien showed in 1893 that the energy dependence is given by

$$u(\omega) = \omega^3 f(\frac{\omega}{T}), \tag{211}$$

where $\omega = 2\pi\nu$ and f is a function that only depends on the ratio ω/T, where T is the temperature. Inside the cavity the radiation oscillates between the walls. In the x direction, the electromagnetic modes are determined by $\sin k_x x$ which has to be zero at $x = 0$ and $x = L$, i.e., the sides of the cavity. This means that $k_x L = n_x \pi$, with $n_x = 1, 2, 3, \ldots$. This implies that

$$(k_x^2 + k_y^2 + k_z^2)L^2 = k^2 L^2 = \left(\frac{\omega L}{c}\right)^2 = (n_x^2 + n_y^2 + n_z^2)\pi^2. \tag{212}$$

Now suppose we make a small adiabatic change of L. This does not effect $n_x^2 + n_y^2 + n_z^2$, since these are just integers. This means that ωL has to be constant or $\omega^3 V$ is constant.

However, we are more interested in a temperature change than a volume change. We can relate the two by making use of the thermodynamic identity

$$TdS = dE + pdV. \tag{213}$$

This says that the change in heat as a result of an entropy change is a result of a change in internal energy (internal forces) plus the work done when changing the volume. This gives for the pressure

$$p = -\left(\frac{\partial E}{\partial V}\right)_T + T\left(\frac{\partial S}{\partial V}\right)_T = -\left(\frac{\partial F}{\partial V}\right)_T, \tag{214}$$

which leads to the obvious definition of the free energy $F = U - TS$. From the differential of the free energy $dF = dE - TdS - SdT = -pdV - SdT$, we can relate entropy and pressure via the Maxwell equations

$$S = -\left(\frac{\partial F}{\partial T}\right)_V \qquad \Rightarrow \qquad \left(\frac{\partial S}{\partial V}\right)_T = -\frac{\partial^2 F}{\partial V \partial T} = \left(\frac{\partial p}{\partial T}\right)_V \tag{215}$$

The pressure is then given by

$$p = -\left(\frac{\partial E}{\partial V}\right)_T + T\left(\frac{\partial p}{\partial T}\right)_V. \tag{216}$$

The internal energy of the particles causes the pressure on the walls of the cavity. A particle reflecting on the wall in the x direction gives a momentum transfer of $2mv_x$. It then takes a time $T_{\text{period}} = 2L/v_x$ to go to the other wall and back again. This leads to a pressure of $p = N2mv_x/T_{\text{period}}L^2 = nmv_x^2$ with $n = N/L^3$ the particle density. Since the cavity is equivalent in all directions $v_x^2 = v_y^2 = v_z^2 = \frac{1}{3}v^2$. So we can write the pressure as $p = \frac{1}{3}nmv^2$. For an ideal gas, we can rewrite this as $p = \frac{2}{3}n(\frac{1}{2}mv^2) = \frac{2}{3}u = \frac{2}{3}U/V$. For a photon gas, however, we have $p = \frac{1}{3}nmc^2 = \frac{1}{3}u$. Writing Eqn. (216) in terms of the internal energy gives

$$\frac{1}{3}u(T) = -\left(\frac{\partial}{\partial V}(u(T)V)\right)_T + T\frac{1}{3}u'(T) = -u(T) + T\frac{1}{3}u'(T) \quad \Rightarrow \quad \frac{1}{3}u'(T) = \frac{4}{3}\frac{u(T)}{T} \quad \Rightarrow \quad u(T) = aT^4, \tag{217}$$

where a is a constant. This is known as the Stefan-Boltzmann law. Therefore the pressure is given by $p = \frac{1}{3}aT^4$. This can be related to the entropy by

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V = \frac{4}{3}aT^3 \qquad \Rightarrow \qquad S = \frac{4}{3}aT^3V = \frac{4}{3}a'\left(\frac{T}{\omega}\right)^3, \tag{218}$$

since $\omega^3 V =$ constant. For an adiabatic transformation of the system, the entropy S is constant, this means that T/ω has to be constant. This can also be written as $T\lambda =$ constant, which is known as Wien's displacement law. Since the total entropy only depends on ω/T, this implies that each oscillator also only depends on ω/T. The total energy can then be found by summing of all the oscillator frequencies

$$u = \sum_\omega u_\omega(T) = \sum_\omega \frac{3}{4}Ts_\omega(\omega/T) = \sum_\omega \omega f_0(\omega/T), \tag{219}$$

using the fact that $u = aT^4 = \frac{3}{4}Ts$. We like to replace the summation by an integral. For that we need to know the density of oscillators. This can be found from Eqn. (212),

$$\frac{\omega^2 L^2}{c^2} = \pi^2 n^2 \quad \Rightarrow \quad n = \frac{\omega L}{c\pi} \quad \Rightarrow \quad 2\frac{1}{8}4\pi n^2 dn = \pi\left(\frac{L}{c\pi}\right)^3 \omega^2 d\omega = Vg(\omega)d\omega \quad \Rightarrow \quad g(\omega) = \frac{\omega^2}{c^3\pi^2} \tag{220}$$

with $n^2 = n_x^2 + n_y^2 + n_z^2$. The factor 2 comes from the two different polarization vectors of the light. The factor $\frac{1}{8}$ comes from taking only the positive k values. Therefore

$$u = \sum_\omega u_\omega(T) = \int d\omega u(\omega) \qquad \Rightarrow \qquad u(\omega) = \omega f_0(\frac{\omega}{T})g(\omega) = \omega^3 f(\frac{\omega}{T}). \tag{221}$$

This is known as Wien's law. To describe the energy dependence Wien suggested that $u(\omega) = \omega^3 \exp(-a\omega/T)$, in analogy with the Boltzmann factor. This gives a remarkably good agreement with the experimentally observed data. If it was not for very recent data that showed ω^2 instead of a ω^3 dependence at low ω, Planck might never have started looking for a solution to this problem.

The low frequency behavior can actually be derived classically. The electromagnetic waves bounding between the walls behave like oscillators. A classical oscillator is described by the equation of motion

$$m\frac{d^2x}{dt^2} = -Kx \qquad \Rightarrow \qquad x = x_0 \sin\omega t \qquad \Rightarrow \qquad -m\omega^2 x = -Kx \qquad \Rightarrow \qquad \omega^2 = \frac{K}{m}. \tag{222}$$

The total energy is thus

$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}Kx^2 = \frac{1}{2}m\omega^2 x_0^2 \cos^2\omega t + \frac{1}{2}m\omega^2 x_0^2 \sin^2\omega t = \frac{1}{2}m\omega^2 x_0^2. \tag{223}$$

The energy can thus be continuously changed depending on the frequency and amplitude of the oscillation. The probability of finding an oscillator with a certain energy is determined by the Boltzmann factor $\exp(-\beta E)$ with $\beta = 1/k_B T$. The average energy can be calculated by integrating the energies multiplied by the probability of finding that energy, and in the end dividing by the total probability:

$$Z = \int dE e^{-\beta E} = -\frac{1}{\beta}\left[e^{-\beta E}\right]_0^\infty = \frac{1}{\beta} = k_B T \tag{224}$$

The average energy can then be calculated

$$\overline{E} = \frac{1}{Z}\int dE E e^{-\beta E} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\beta\frac{\partial}{\partial\beta}\left(\frac{1}{\beta}\right) = \frac{1}{\beta} = k_B T. \tag{225}$$

Note that there is no energy dependence. This is a typical classical equipartition effect. Each oscillator has the same energy. Of course when ω is, large the amplitude of the oscillation will be small. The Rayleigh-Jeans law can now be derived as

$$u(\omega) = g(\omega)\overline{E} = \frac{\omega^2}{c^3\pi^2}k_BT. \tag{226}$$

(where the significant contribution of Jeans was the introduction of the factor $\frac{1}{8}$ in our Eqn (220)). Note that the Rayleigh-Jeans equation has a temperature dependence but just increases as a function of ω. Integrating over ω would give an infinite internal energy in our cavity. This was later called the ultraviolet catastrophe.

At this point Planck enters the scene. Planck worked for most of his life on entropy so that is where he started. Let us take the Rayleigh-Jeans law. From the definition of temperature, we get

$$\left(\frac{\partial S}{\partial \overline{E}}\right)_V = \frac{1}{T} = \frac{k_B}{\overline{E}}. \tag{227}$$

Differentiating this to $\overline{E}$,

$$\left(\frac{\partial^2 S}{\partial \overline{E}^2}\right)_V = -\frac{k_B}{\overline{E}^2}. \tag{228}$$

Note that this is negative since the entropy is maximum. For Wien's law, we obtain

$$\overline{E} = \text{const} \times \omega e^{-\alpha\frac{\omega}{T}} \qquad \Rightarrow \qquad T = \frac{\text{const} \times \omega}{\ln\frac{\overline{E}}{\omega}}, \tag{229}$$

where the constants are not necessarily equal. This gives

$$\left(\frac{\partial S}{\partial \overline{E}}\right)_V = \frac{1}{T} = \text{const} \times \frac{\ln\frac{\overline{E}}{\omega}}{\omega} \qquad \Rightarrow \qquad \left(\frac{\partial^2 S}{\partial \overline{E}^2}\right)_V = \frac{\text{const}}{\overline{E}}. \tag{230}$$

Planck's solution is a simple interpolation between the two results

$$\left(\frac{\partial^2 S}{\partial \overline{E}^2}\right)_V = \frac{-k_B}{\overline{E}^2 + \alpha\overline{E}}. \tag{231}$$

Integrating this gives

$$\left(\frac{\partial S}{\partial \overline{E}}\right)_V = \frac{k_B}{\alpha}\ln(1 + \frac{\alpha}{\overline{E}}) = \frac{1}{T} \qquad \Rightarrow \qquad \overline{E} = \frac{\alpha}{\exp(\frac{\alpha}{k_BT}) - 1}. \tag{232}$$

For the frequency dependence of the blackbody radiation, we then obtain

$$u(\omega) = \frac{\omega^2}{c^3\pi^2}\frac{\alpha}{\exp(\frac{\alpha}{k_BT}) - 1} = \frac{\omega^2}{c^3\pi^2}\frac{\hbar\omega}{\exp(\frac{\hbar\omega}{k_BT}) - 1}, \tag{233}$$

where the factor $\alpha = \text{constant} \times \omega$, since $u(\omega)$ has to satisfy Wien's law. In the equation, we have written $\hbar$ for the constant.

This can also be derived in a somewhat different (more modern) way. If we take the total probability (also known as the partition function) in Eqn. (224) and now sum over the energies instead of integrate

$$Z = \sum_{n=0}^{\infty}\exp(-\beta n\hbar\omega) = \frac{1}{1 - \exp(-\beta\hbar\omega)}. \tag{234}$$

The average energy is then

$$\overline{E} = -\frac{1}{Z}\frac{\partial Z}{\partial\beta} = -\frac{\partial \ln Z}{\partial\beta} = \frac{\partial}{\partial\beta}\ln(1 - \exp(-\beta\hbar\omega)) = \frac{\hbar\omega}{1 - \exp(-\beta\hbar\omega)}\exp(-\beta\hbar\omega) = \frac{\hbar\omega}{\exp(\beta\hbar\omega) - 1}. \tag{235}$$

Or, as Planck did it, suppose we have N oscillators we an average energy $\overline{E}$. The total energy is then $E_{\text{tot}} = N\overline{E}$. Suppose we can only divide the energies among the oscillators in portions of $\hbar\omega$. This means that we have $P = E_{\text{tot}}/\hbar\omega$

portions. Dividing P portions among oscillators goes like $\cdot|\cdots|\cdot\cdot|\cdots$, meaning that we have P dots and we need $N-1$ lines to divide it among the N oscillators. The chance for a specific configuration is

$$W = \frac{(N+P-1)!}{(N-1)!P!}. \tag{236}$$

The entropy is defined as

$$S_N = k_B \ln W \cong k_B \left[(N+P)\ln(N+P) - N\ln N - P\ln P\right], \tag{237}$$

where we have assumed that N and P are large so we can apply Stirling's formula $N! \sim N^N/e^N$. The entropy per oscillator is then

$$S = \frac{S_N}{N} = k_B(1+\frac{P}{N})\ln(1+\frac{P}{N}) - k_B\frac{P}{N}\ln\frac{P}{N} = k_B(1+\frac{\overline{E}}{\hbar\omega})\ln(1+\frac{\overline{E}}{\hbar\omega}) - k_B\frac{\overline{E}}{\hbar\omega}\ln\frac{\overline{E}}{\hbar\omega}. \tag{238}$$

Taking the derivative

$$\frac{dS}{d\overline{E}} = k_B\frac{1}{\hbar\omega}\ln(1+\frac{\overline{E}}{\hbar\omega}) + \frac{k_B}{\hbar\omega} - k_B\frac{1}{\hbar\omega}\ln\frac{\overline{E}}{\hbar\omega} - \frac{k_B}{\hbar\omega} \tag{239}$$

and differentiating again

$$\frac{d^2S}{d\overline{E}^2} = \frac{k_B}{\hbar\omega}\left(\frac{1}{\overline{E}+\hbar\omega} - \frac{1}{\overline{E}}\right) = \frac{-k_B}{(\overline{E}+\hbar\omega)\overline{E}}, \tag{240}$$

which rederives Planck's interpolation formula.

B. Many-particle wavefunctions

See Reif Chapter 9, lecture notes MvV 560
In quantum mechanics, particles are indistinguishable. This means that we should be able to interchange two particles without seeing any effect on the probability function. We can distinguish two fundamentally different types of particles, namely those with a symmetric and antisymmetric wavefunction. Under permutation between particle i and j for a N particle wavefunction, one has the following possibilities

$$P_{ij}\psi^{\pm}_{k_1\cdots k_N}(\mathbf{r}_1,\cdots,\mathbf{r}_i,\cdots,\mathbf{r}_j,\cdots,\mathbf{r}_N) = \pm\psi^{\pm}_{k_1\cdots k_N}(\mathbf{r}_1,\cdots,\mathbf{r}_i,\cdots,\mathbf{r}_j,\cdots,\mathbf{r}_N) \tag{241}$$

where k_i means that a particle is in a certain quantum state n_i (for example a particle in a box where the quantum numbers are $n_i = n_x, n_y, n_z$, but it could also refer to the quantum numbers of an atom $n_i = nlm\sigma$). At first sight one might think it does not make any difference since the real properties are related to the square of the wavefunction and a plus or minus sign should not make too much of a difference. However, the sign has drastic consequences for the statistics. Let us look at the wavefunction for two particles:

$$\psi^{\pm}_{k_1,k_2}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}(\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2) \pm \varphi_{k_2}(\mathbf{r}_1)\varphi_{k_1}(\mathbf{r}_2)). \tag{242}$$

where ϕ_k are eigenfunctions of the Hamiltonian. The symmetric and antisymmetric wavefunctions give two fundamentally different particles. Particles whose wavefunction is symmetric are bosons, particles with antisymmetric wavefunctions are called fermions. Examples of bosons are photons, π and K mesons. These are particles with integral spin. Also the harmonic oscillator and lattice vibrations (also known as phonons) behave like a "boson"-like particle. Particles with half integral spin behave like fermions. Examples are electrons, protons, and neutrons. The antisymmetric wavefunctions have a very special characteristic. Let us take two particle in the same quantum state

$$\psi^{-}_{k,k}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}(\varphi_k(\mathbf{r}_1)\varphi_k(\mathbf{r}_2) - \varphi_k(\mathbf{r}_1)\varphi_k(\mathbf{r}_2)) = 0. \tag{243}$$

Therefore, it is impossible to have two fermions with the same quantum numbers. Note that the spin is also considered a quantum number. Therefore, we can have two electrons in the same wavefunction (say the same plane wave or the same atomic orbital), but with opposite spins. The occupation of each fermion state is therefore 0 or 1.

Let us now generalize the wavefunctions for an arbitrary number of particles. For bosons the general wavefunction is

$$\Psi_{k_1,\ldots,k_N} = \sqrt{\frac{\Pi_i n_{k_i}!}{N!}} \sum_P \psi_{k_1}(\mathbf{r}_1)\cdots\psi_{k_N}(\mathbf{r}_N), \tag{244}$$

where the summations runs over all the possible permutations of *different* indices k_i. This gives the normalization constant. n_{k_i} indicates how many k-values are k_i. Note that the $\sum_i n_{k_i} = N$. For two particles with $S = 0$, we have

$$\Psi_{k_1,k_2} = \frac{1}{\sqrt{2}} \left\{\psi_{k_1}(\mathbf{r}_1)\psi_{k_2}(\mathbf{r}_2) + \psi_{k_2}(\mathbf{r}_1)\psi_{k_1}(\mathbf{r}_2)\right\}. \tag{245}$$

For $k_1 = k_2 = k$, we find

$$\Psi_{k_1,k_2} = \sqrt{\frac{2!}{2!}}\psi_k(\mathbf{r}_1)\psi_k(\mathbf{r}_2) = \psi_k(\mathbf{r}_1)\psi_k(\mathbf{r}_2), \tag{246}$$

using the fact that $n_k = 2$. We see that it is possible to put two bosons in the same quantum state. In fact, we can put any arbitrary number of bosons in a particle quantum state. From a statistical point of view, we therefore have for the occupation numbers of a particular quantum state i

$$n_i = 0, 1 \quad \text{fermions} \tag{247}$$
$$n_i = 0, 1, 2, 3, \cdots \quad \text{bosons} \tag{248}$$

This is an important aspect that we have to include in our quantum statistics. Up till now the only aspect that we have used is the indistinguishability.

The situation for electrons (and other spin-1/2 particles) is fact a bit more complicated since the total wavefunction, i.e., orbital and spin, has to be asymmetric. This means $\psi_S(\mathbf{r}_1, \mathbf{r}_2)\xi_A(1, 2)$ or $\psi_A(\mathbf{r}_1, \mathbf{r}_2)\xi_S(1, 2)$, where $\xi(1, 2)$ is the wavefunction for the spin. There are four ways to combine two spins $\xi_\uparrow(1)\xi_\uparrow(2)$, $\xi_\uparrow(1)\xi_\downarrow(2)$, $\xi_\downarrow(1)\xi_\uparrow(2)$, and $\xi_\downarrow(1)\xi_\downarrow(2)$. $\xi_\uparrow(1)\xi_\uparrow(2)$ and $\xi_\downarrow(1)\xi_\downarrow(2)$ are clearly symmetric spin functions, since interchange of the particles does not change the sign of the wave function. These are the $\xi_{11}(1, 2)$ and $\xi_{1,-1}(1, 2)$ components of the triplet ($S = 1$). Since $S_z = 1, 0, -1$ for $S = 1$, there should also be a $S = 0$ component. This can be obtained by using the step operator for the spin

$$S_-\xi_{11} = \sqrt{(1+1)(1-1+1)}\xi_{10} = \sqrt{2}\xi_{10} = \xi_\uparrow(1)\xi_\downarrow(2) + \xi_\downarrow(1)\xi_\uparrow(2) \tag{249}$$
$$\Rightarrow \qquad \xi_{10} = \frac{1}{\sqrt{2}}\left[\xi_\uparrow(1)\xi_\downarrow(2) + \xi_\downarrow(1)\xi_\uparrow(2)\right] \tag{250}$$

where we have used

$$S_\pm\xi_{SS_z} = \sqrt{(S \mp S_z)(S \pm S_z + 1)}\xi_{S,S_z\pm1} \tag{251}$$

Note that this is also a symmetric wavefunction. The remaining wave function has to be orthogonal to the previous one

$$\xi_{00}(1, 2) = \frac{1}{\sqrt{2}}\left[\xi_\uparrow(1)\xi_\downarrow(2) - \xi_\downarrow(1)\xi_\uparrow(2)\right] \tag{252}$$

which is antisymmetric. Since there is only one component (called a singlet), this has to be $S = 0$. The four total wavefunctions for two electrons are therefore $\Psi_{1S_z}(\mathbf{r}_1, \mathbf{r}_2) = \psi_A(\mathbf{r}_1, \mathbf{r}_2)\xi_{1S_z}(1, 2)$ with $S_z = 1, 0, -1$ and $\Psi_{00}(\mathbf{r}_1, \mathbf{r}_2) = \psi_S(\mathbf{r}_1, \mathbf{r}_2)\xi_{00}(1, 2)$. The $S = 1$ and $S_z = 1$ is given by

$$\Psi_{11}(\mathbf{r}_1, \mathbf{r}_2) = \psi_A(\mathbf{r}_1, \mathbf{r}_2)\xi_{11}(1, 2) = \frac{1}{\sqrt{2}}\left[\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2) - \varphi_{k_2}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_1)\right]\xi_\uparrow(1)\xi_\uparrow(2) \tag{253}$$
$$= \frac{1}{\sqrt{2}}\left[\varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2) - \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(2)\right] \tag{254}$$
$$= \frac{1}{\sqrt{2}}\begin{vmatrix} \varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_1}(\mathbf{r}_2)\xi_\uparrow(2) \\ \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2) \end{vmatrix} \tag{255}$$

The $S=0$ and $S_z=0$ wavefunction is given by

$$\Psi_{00}(\mathbf{r}_1,\mathbf{r}_2)=\psi_S(\mathbf{r}_1,\mathbf{r}_2)\xi_{00}(1,2) = \frac{1}{\sqrt{2}}\left[\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2)+\varphi_{k_2}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_1)\right]\frac{1}{\sqrt{2}}\left[\xi_\uparrow(1)\xi_\downarrow(2)-\xi_\downarrow(1)\xi_\uparrow(2)\right] \quad (256)$$

$$= \frac{1}{2}\left[\varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_2}(\mathbf{r}_2)\xi_\downarrow(2)-\varphi_{k_2}(\mathbf{r}_1)\xi_\downarrow(1)\varphi_{k_1}(\mathbf{r}_2)\xi_\uparrow(2)\right. \quad (257)$$

$$\left.-\varphi_{k_1}(\mathbf{r}_1)\xi_\downarrow(1)\varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2)+\varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_1}(\mathbf{r}_2)\xi_\downarrow(2)\right] \quad (258)$$

Note that this can also be written as a combination of Slater determinants

$$\Psi_{00}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{vmatrix}\varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_1}(\mathbf{r}_2)\xi_\uparrow(2)\\ \varphi_{k_2}(\mathbf{r}_1)\xi_\downarrow(1) & \varphi_{k_2}(\mathbf{r}_2)\xi_\downarrow(2)\end{vmatrix}-\frac{1}{\sqrt{2}}\begin{vmatrix}\varphi_{k_1}(\mathbf{r}_1)\xi_\downarrow(1) & \varphi_{k_1}(\mathbf{r}_2)\xi_\downarrow(2)\\ \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2)\end{vmatrix}\right) \quad (259)$$

$$= \frac{1}{\sqrt{2}}[\psi_{k_1\uparrow,k_2\downarrow}(\mathbf{r}_1,\mathbf{r}_2)-\psi_{k_1\downarrow,k_2\uparrow}(\mathbf{r}_1,\mathbf{r}_2)] \quad (260)$$

Note that in general we can built up antisymmetric states from Slater determinants

$$\psi_{k_1\sigma_1,\ldots,k_N\sigma_N}(\mathbf{r}_1\cdots\mathbf{r}_N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\varphi_{k_1}(\mathbf{r}_1)\xi_{\sigma_1}(1) & \cdots & \varphi_{k_1}(\mathbf{r}_N)\xi_{\sigma_1}(N)\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \varphi_{k_N}(\mathbf{r}_1)\xi_{\sigma_N}(1) & \cdots & \varphi_{k_N}(\mathbf{r}_N)\xi_{\sigma_N}(N)\end{vmatrix} \quad (261)$$

where $\sigma_i=\uparrow,\downarrow$. As a result of the properties of the determinant, having two columns with the same k and σ implies that the determinant is zero, just as for the exclusion principle. Interchange of two columns also leads to a change in sign of the determinant, giving us directly the antisymmetry. This is obvious for the $S_z=1,-1$ states of the triplet. For the $S_z=0$ states, one needs a combination of Slater determinants

$$\Psi_{k_1k_2,00}=\frac{1}{\sqrt{2}}\left(\psi_{k_1\uparrow,k_2\downarrow}-\psi_{k_1\downarrow,k_2\uparrow}\right) \qquad \text{and} \qquad \Psi_{k_1k_2,10}=\frac{1}{\sqrt{2}}\left(\psi_{k_1\uparrow,k_2\downarrow}+\psi_{k_1\downarrow,k_2\uparrow}\right). \quad (262)$$

You may have noticed that dealing with the statistic of many particles using wavefunction is quite cumbersome. This is due to the fact that we have to permute all the different particles (or the different coordinates $\mathbf{r}_i$ over all the different wavefunctions. The statistics for fermions and bosons can greatly simplify if we start out with something indistinguishable. This can be done by using operators. You probably remember that the analytical solution of the harmonic oscillator involved the step up/down operators $a^\dagger$ and a, which had the relations

$$a^\dagger|n\rangle=\sqrt{n+1}|n+1\rangle \quad \text{and} \quad a|n\rangle=\sqrt{n}|n-1\rangle. \quad (263)$$

Now forget for the moment the relation of $a^\dagger$ with x and p, since this result is far more general. Note that $a^\dagger$ effectively put adds one to the quantum state, whereas a removes one from the quantum state. Let us now identify this with bosons. $a^\dagger$ creates a boson in the quantum state, whereas a removes a boson from the quantum state. Now the harmonic oscillator is only one quantum state. We can in principle have many different quantum states k in which we can put bosons. So we can generalize it to

$$a_k^\dagger|n_1,n_2,\cdots,n_k,\cdots\rangle = \sqrt{n_k+1}|n_1,n_2,\cdots,n_k+1,\cdots\rangle \quad \text{and} \quad (264)$$

$$a_k|n_1,n_2,\cdots,n_k,\cdots\rangle = \sqrt{n_k}|n_1,n_2,\cdots,n_k-1,\cdots\rangle. \quad (265)$$

These operators have some nice characteristics: they are indistinguishable. In addition, they are symmetric. This follows from the generalized commutation relations

$$[a_k,a_{k'}^\dagger]=\delta_{k,k'} \qquad \text{and} \qquad [a_k^\dagger,a_{k'}^\dagger]=0 \qquad \text{and} \qquad [a_k,a_k']=0 \quad (266)$$

Thus for $k\neq k'$, we can write

$$a_k^\dagger a_{k'}^\dagger=a_{k'}^\dagger a_k^\dagger. \quad (267)$$

Therefore, interchange of two particles does not change the sign of the wavefunction.

The wavefunction for a single harmonic oscillator could be written as

$$|n\rangle=\frac{1}{\sqrt{N!}}(a^\dagger)^N|0\rangle. \quad (268)$$

For more than one k value, we can generalize the wavefunction as

$$|\{n_{k_i}\}\rangle = \frac{1}{\sqrt{\Pi_i n_{k_i}}}\Pi_i (a^\dagger_{k_i})^{n_{k_i}}|0\rangle. \tag{269}$$

Let us look at the two-particle wavefunctions. If the two particles are in the same state (i.e. have the same quantum number) then the wavefunction should be similar to that of the harmonic oscillator:

$$|kk\rangle = \frac{1}{\sqrt{2}} a^\dagger_k a^\dagger_k |0\rangle. \tag{270}$$

This wavefunction is now properly normalized since

$$\begin{aligned}
\langle kk|kk\rangle &= \frac{1}{2}\langle 0|a_k a_k a^\dagger_k a^\dagger_k|0\rangle = \frac{1}{2}\langle 0|a_k(1+a^\dagger_k a_k)a^\dagger_k|0\rangle && (271)\\
&= \frac{1}{2}\langle 0|a_k a^\dagger_k|0\rangle + \frac{1}{2}\langle 0|a_k a^\dagger_k a_k a^\dagger_k|0\rangle && (272)\\
&= \frac{1}{2}\langle 0|(1+a^\dagger_k a_k)|0\rangle + \frac{1}{2}\langle 0|(1+a^\dagger_k a_k)(1+a^\dagger_k a_k)|0\rangle = 1 && (273)\\
& && (274)
\end{aligned}$$

since $a_k|0\rangle = 0$. For two different wavevectors, we have

$$|k_1 k_2\rangle = a^\dagger_{k_1} a^\dagger_{k_2}|0\rangle. \tag{275}$$

This gives the right normalization since

$$\langle k_1 k_2|k_1 k_2\rangle = \langle 0|a_{k_1}a_{k_2}a^\dagger_{k_2}a^\dagger_{k_1}|0\rangle = \langle 0|a_{k_1}(1+a^\dagger_{k_2}a_{k_2})a^\dagger_{k_1}|0\rangle = \langle 0|a_{k_1}a^\dagger_{k_1}|0\rangle = \langle 0|(1+a^\dagger_{k_1}a_{k_1})|0\rangle = 1 \tag{276}$$

Working with Slater determinants becomes quite elaborate. We can however do the same trick with operators to account for the statistics as with bosons. The operator $c^\dagger_k$ creates a fermion in a certain quantum state k, whereas c_k removes it from this quantum state, i.e.

$$c^\dagger_k|0_k\rangle = |1_k\rangle \quad \text{and} \quad c_k|1_k\rangle = |0_k\rangle, \tag{277}$$

where $|0_k\rangle$ and $|1_k\rangle$ denote the states with 0 and 1 electron in quantum state k, respectively. However, unlike bosons that satisfy commutation relations, fermions satisfy anticommutation relations

$$\{c_k, c^\dagger_{k'}\} = \delta_{k,k'}, \qquad \{c^\dagger_k, c^\dagger_{k'}\} = 0, \qquad \{c_k, c_{k'}\} = 0, \tag{278}$$

where the term in braces is defined as

$$\{A, B\} = AB + BA. \tag{279}$$

However, we cannot put two particles in the same state, since

$$c^\dagger_k c^\dagger_k|0\rangle = \frac{1}{2}\{c^\dagger_k, c^\dagger_k\}|0\rangle = 0. \tag{280}$$

We can put them in different quantum states, thus for $k \neq k'$

$$c^\dagger_k c^\dagger_{k'}|0\rangle = -c^\dagger_{k'} c^\dagger_k|0\rangle. \tag{281}$$

Note that the operators directly take care of the antisymmetry of the wavefunctions. Including spin, we can write the wavefunctions $|k_1 k_2, SS_z\rangle$ for two particles in states labeled with quantum numbers k_1 and k_2 with different spin S and S_z as

$$|k_1 k_2, 11\rangle = c^\dagger_{k_2\uparrow} c^\dagger_{k_1\uparrow}|0\rangle \quad , \quad |k_1 k_2, 10\rangle = \frac{1}{\sqrt{2}}\left(c^\dagger_{k_2\downarrow} c^\dagger_{k_1\uparrow} + c^\dagger_{k_2\uparrow} c^\dagger_{k_1\downarrow}\right)|0\rangle \quad , \quad |k_1 k_2, 1, -1\rangle = c^\dagger_{k_2\downarrow} c^\dagger_{k_1\downarrow}|0\rangle \tag{282}$$

$$|k_1 k_2, 00\rangle = \frac{1}{\sqrt{2}}\left(c^\dagger_{k_2\downarrow} c^\dagger_{k_1\uparrow} - c^\dagger_{k_2\uparrow} c^\dagger_{k_1\downarrow}\right)|0\rangle. \tag{283}$$

In general, we can write the many-body Fermion wavefunction as

$$|\{n_{k_i}\}\rangle = \Pi_i (c^\dagger_{k_i})|0\rangle. \tag{284}$$

Note, that the normalization is unity since we can never have more than one fermion in the same quantum state.

Summarizing, we have seen that the a symmetric and antisymmetric wavefunction corresponds to particles with completely different statistics. Bosons have a symmetric wavefunction that allows any number of bosons in a particular quantum state. Fermions have antisymmetric wavefunctions and the occupation numbers are restricted to 0 and 1. We also saw that these statistic can be closely related to operators and their corresponding commutation relationships. Bosons commute, whereas fermions anticommute.

C. The chemical potential

Before discussing the different quantum distribution functions, it is convenient to introduce the concept of chemical potential. So far, we have looked at systems that are in thermal contact with each other, meaning that they can exchange energy. However, let us now consider that the two systems can also exchange particles with each other. Let us take system 1 with volume V_1 and particle number N_1 and system 2 with volume V_2 and particle number N_2. Both systems are at a temperature. The total free energy is given by the sum of the two systems

$$F = F_1 + F_2 = U_1 + U_2 - T(S_1 + S_2). \tag{285}$$

The systems can now interchange particles with each other. For the lowest free energy, the differential dF,

$$dF = \left(\frac{\partial F}{\partial N_1}\right)_{T,V_1} dN_1 + \left(\frac{\partial F}{\partial N_2}\right)_{T,V_2} dN_2 \tag{286}$$

equals zero. Since $dN_2 = -dN_1$, we have

$$\left(\frac{\partial F}{\partial N_1}\right)_{T,V_1} dN_1 - \left(\frac{\partial F}{\partial N_2}\right)_{T,V_2} dN_1 = 0, \tag{287}$$

or

$$\left(\frac{\partial F}{\partial N_1}\right)_{T,V_1} = \left(\frac{\partial F}{\partial N_2}\right)_{T,V_2}. \tag{288}$$

The two system are in equilibrium when the two partial derivatives are equal. Let us identify this as the chemical potential

$$\mu_B \equiv \left(\frac{\partial F}{\partial N}\right)_{T,V}. \tag{289}$$

Equilibrium then occurs when the chemical potential are equal $\mu_1 = \mu_2$. Since the number of particles is discrete this should in principle be defined as

$$\mu \equiv F(T,V,N) - F(T,V,N-1) \tag{290}$$

and therefore corresponds to the energy gained when a particle is added to the system.

D. Quantum distribution functions

Reif Chapter 9 We want to evaluate the average number of particles for different quantum statistics. As usual we start with the partition function

$$Z = \sum_{\{n_i\}} e^{-\beta(n_1\varepsilon_1 + n_2\varepsilon_2 + \cdots)}, \tag{291}$$

where the summation $\{n_i\}$ goes over all the possible permutations of occupation numbers. We now want to calculate the mean occupation in a particular quantum state n_s,

$$\overline{n}_s = \frac{\sum_{\{n_i\}} n_s e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots)}}{\sum_{\{n_i\}} e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots)}}. \tag{292}$$

We can extract n_s from the exponential by taking the derivative with respect to ε_s at fixed temperature and volume

$$\overline{n}_s = \frac{1}{Z}\sum_{\{n_i\}}\left(-\frac{1}{\beta}\frac{\partial}{\partial\varepsilon_s}\right)e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots)} \tag{293}$$

$$= -\frac{1}{\beta}\frac{1}{Z}\frac{\partial Z}{\partial\varepsilon_s} = -\frac{1}{\beta}\frac{\partial \ln Z}{\partial\varepsilon_s} = -k_BT\frac{\partial \ln Z}{\partial\varepsilon_s} = \left(\frac{\partial F}{\partial\varepsilon_s}\right)_{T,V}. \tag{294}$$

1. Photon statistics

Photon statistics is related to Bose-Einstein statistics, i.e. the statistics that applies to bosons, with an unspecified number of particles. This occurs for example for photons and phonons (lattice vibrations). Let us separate the terms including the state s

$$\overline{n}_s = \frac{\sum_{n_s} n_s e^{-\beta n_s\varepsilon_s}}{\sum_{n_s} e^{-\beta n_s\varepsilon_s}}\frac{\sum_{\{n_i\},i\neq s} e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots)}}{\sum_{\{n_i\},i\neq s} e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots)}}. \tag{295}$$

Since there are no restriction to the number of particles the sum excluding the quantum state s are equal and cancel. We are left with

$$\overline{n}_s = \frac{\sum_{n_s} n_s e^{-\beta n_s\varepsilon_s}}{\sum_{n_s} e^{-\beta n_s\varepsilon_s}} = -\frac{1}{\beta}\frac{\partial}{\partial\varepsilon_s}\ln\left(\sum_{n_s} e^{-\beta n_s\varepsilon_s}\right). \tag{296}$$

The summation is a simple geometric series

$$\sum_{n_s} e^{-\beta n_s\varepsilon_s} = 1 + e^{-\beta\varepsilon_s} + e^{-2\beta\varepsilon_s} + \cdots = \frac{1}{1-e^{-\beta\varepsilon_s}} \tag{297}$$

This gives

$$\overline{n}_s = -\frac{1}{\beta}\frac{\partial}{\partial\varepsilon_s}\ln\frac{1}{1-e^{-\beta\varepsilon_s}} = \frac{1}{\beta}\frac{\partial}{\partial\varepsilon_s}\ln(1-e^{-\beta\varepsilon_s}) = \frac{e^{-\beta\varepsilon_s}}{1-e^{-\beta\varepsilon_s}} \tag{298}$$

or

$$\overline{n}_s = \frac{1}{e^{\beta\varepsilon_s}-1} \tag{299}$$

This is the well-known Planck distribution.

2. Fermi-Dirac statistics

Let us now turn our attention to systems where the number of particles is fixed. We therefore need to satisfy that

$$N = \sum_s n_s. \tag{300}$$

We want to evaluate the sum over the states excluding s in Eqn. (295)

$$Z_s(N) = \sum_{\{n_i\},i\neq s} e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots)} \tag{301}$$

In the state s, we can either put 0 or 1 particles. The average number of particles in n_s then becomes

$$\overline{n}_s = \frac{0 \times e^0 \times Z_s(N) + 1 \times e^{-\beta\varepsilon_s} Z_s(N-1)}{e^0 \times Z_s(N) + e^{-\beta\varepsilon_s} Z_s(N-1)} = \frac{e^{-\beta\varepsilon_s} Z_s(N-1)}{Z_s(N) + e^{-\beta\varepsilon_s} Z_s(N-1)}, \tag{302}$$

or

$$\overline{n}_s = \frac{1}{\frac{Z_s(N)}{Z_s(N-1)} e^{\beta\varepsilon_s} + 1}. \tag{303}$$

For the evaluation of the change in $Z_s(N)$ as a function of N, we make use of its relation to the free energy

$$\ln Z_s(N - dN) = \ln Z_s(N) - \left(\frac{\partial \ln Z(N)}{\partial N}\right)_{T,V} dN = \ln Z_s(N) + \beta \left(\frac{\partial F}{\partial N}\right)_{T,V} dN. \tag{304}$$

However, the latter partial derivative is by definition the chemical potential, see Eqn. (289).

$$\ln Z_s(N - dN) = \ln Z_s(N) + \beta\mu dN \quad \Rightarrow \quad Z_s(N - dN) = Z_s(N) e^{\beta\mu dN} \tag{305}$$

or

$$Z_s(N-1) = Z_s(N) e^{\beta\mu}. \tag{306}$$

Inserting this into the expression for $\overline{n}_s$ gives

$$\overline{n}_s = \frac{1}{e^{\beta(\varepsilon_s - \mu)} + 1}. \tag{307}$$

This expression is known as the Fermi-Dirac function. Since the exponent is always greater than one, $\overline{n}_s \leq 1$. When $\varepsilon_s \to \infty$, $\overline{n}_s \to 0$. Note that $\overline{n}_s = \frac{1}{2}$, when $\varepsilon_s = \mu$.

3. Bose-Einstein statistics

The approach for bosonic particles is similar to that for Fermi-Dirac statistics, except that the allowed quantum states are now $n_s = 0, 1, 2, 3, \cdots$.

$$\overline{n}_s = \frac{0 \times e^0 \times Z_s(N) + 1 \times e^{-\beta\varepsilon_s} Z_s(N-1) + 2 \times e^{-2\beta\varepsilon_s} Z_s(N-2) + \cdots}{e^0 \times Z_s(N) + e^{-\beta\varepsilon_s} Z_s(N-1) + e^{-2\beta\varepsilon_s} Z_s(N-2) + \cdots} \tag{308}$$

Using the relationship in Eqn. (305), this becomes

$$\overline{n}_s = \frac{Z_s(N)[0 + 1 \times e^{-\beta(\varepsilon_s - \mu)} + 2 \times e^{-2\beta(\varepsilon_s - \mu)} + \cdots]}{Z_s(N)[1 + e^{-\beta(\varepsilon_s - \mu)} + e^{-2\beta(\varepsilon_s - \mu)} + \cdots]} = \frac{\sum_{n_s} n_s e^{-\beta n_s(\varepsilon_s - \mu)}}{\sum_{n_s} e^{-\beta n_s(\varepsilon_s - \mu)}} \tag{309}$$

But this is, apart from the presence of a chemical potential, equivalent to what we found in Eqn. (296). The Bose-Einstein distribution function is therefore

$$\overline{n}_s = \frac{1}{e^{\beta(\varepsilon_s - \mu)} - 1}. \tag{310}$$

4. Maxwell-Boltzmann statistics

For comparison, it is instructive to look at the case of distinguishable particles, known as Maxwell-Boltzmann statistics. Since the particles are distinguishable, we now have to account for the number of permutations possible between the different particles. For example, particles a and b in state 1, and c in state 2 is now different from particles a and c in state 1, and b in state 2. For quantum statistics, these states are equivalent, namely two particles in state

1 and one in state 2. To pick n_1 particles from N particles can be done in $\begin{pmatrix} N \\ n_1 \end{pmatrix}$ ways. To pick n_2 particles from the remaining $N - n_1$ particles can be done in $\begin{pmatrix} N-n_1 \\ n_1 \end{pmatrix}$ ways. And so on. The total probability is the product

$$\begin{pmatrix} N \\ n_1 \end{pmatrix} \begin{pmatrix} N-n_1 \\ n_2 \end{pmatrix} \begin{pmatrix} N-n_1-n_2 \\ n_3 \end{pmatrix} \cdots = \frac{N!}{n_1!(N-n_1!)} \frac{(N-n_1)!}{n_2!(N-n_1-n_2!)} \frac{(N-n_1-n_2)!}{n_3!(N-n_1-n_2-n_3!)} \cdots = \frac{N!}{n_1!n_2!n_3!\cdots}.$$

The partition function is then

$$Z = \sum_{\{n_i\}} \frac{N!}{n_1!n_2!n_3!\cdots} e^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+n_3\varepsilon_3+\cdots)}. \tag{311}$$

This can also be rewritten as

$$Z = \sum_{\{n_i\}} \frac{N!}{n_1!n_2!n_3!\cdots} (e^{-\beta\varepsilon_1})^{n_1} (e^{-\beta\varepsilon_2})^{n_2} (e^{-\beta\varepsilon_3})^{n_3} \cdots \tag{312}$$

and by the properties of a polynomial as

$$Z = (e^{-\beta\varepsilon_1} + e^{-\beta\varepsilon_2} e^{-\beta\varepsilon_3} + \cdots)^N. \tag{313}$$

Taking the logarithm gives

$$\ln Z = N \ln \left(\sum_i e^{-\beta\varepsilon_i} \right). \tag{314}$$

The occupation of a particular state s is given by

$$\overline{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \varepsilon_s} = -\frac{1}{\beta} N \frac{1}{\sum_i e^{-\beta\varepsilon_i}} \times e^{-\beta\varepsilon_s} \times (-\beta), \tag{315}$$

which gives

$$\overline{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \varepsilon_s} = N \frac{e^{-\beta\varepsilon_s}}{\sum_i e^{-\beta\varepsilon_i}}. \tag{316}$$

However, this is just N times larger that what we would expect classically for the probability of finding a particle in state s using the Boltzmann factor.

V. SYSTEMS OF INTERACTING PARTICLES OR SPINS

A. Van der Waals gas

Stanley, Chapter 5, Pathria Chapter 11
In previous section, we only looked at the ideal gas following the perfect gas law $PV = Nk_BT$. Note that this gas can be cooled down to $T = 0$. Since the motion of the particles is zero, the pressure disappears. Also we can reduce the volume to zero. This leads to an infinite pressure, but we are still dealing with a gas. However, this is in contradiction with experimental observations and our own experiences that if we cool down a gas, it becomes a liquid. Apparently something is missing in our ideal gas model, which after all only contains the kinetic energy of the particles. There is nothing in our model to describe the interaction between the particles or the finite size of the particles (they are basically point charges). Interactions between molecules can be described with a semiempirical potential know as the Lennard-Jones potential,

$$u(r) = u_0 \left[\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right]. \tag{317}$$

This potential is strongly repulsive for small r reflecting the finite size of the molecule. However it is weakly attractive for larger r. Note that this attraction is a result of the polarizability of the molecule, since the molecule itself is

neutral. It is therefore a higher-order effect, which is reflected in the much higher power compared to the usual $1/r$ Coulomb interaction. However, to obtain the Van der Waals equation of state, we can simplify the interactions even further. Let us assume that the essential interactions are a hard core potential of a sphere with radius r_0. The attractive potential is replaced by a constant potential. This bascially means that we neglect all variations in the potential landscape due to the actual position of the molecules. The potential is then

$$u(r) = \begin{cases} \infty & r < r_0 \\ \overline{u} = \frac{\overline{U}}{V} & \text{otherwise} \end{cases} \tag{318}$$

We set $\overline{u} = \frac{\overline{U}}{V}$. The reason for doing is that we do want $\int u(r) d\mathbf{r} = \overline{u}V = \overline{U}$ to be independent of the volume, since this would be the case if the potential was short ranged. Apart from this effective potential describing the interactions between the particles, the particles move independently, i.e. there are no collisions.

The classical partition function for one particle can now be calculated as

$$Z_1^c = \frac{1}{h^3} \int d\mathbf{r} \int d\mathbf{p} \exp[-\beta \frac{p^2}{2m} - \beta U(\mathbf{r})]. \tag{319}$$

Note that $U(\mathbf{r})$ is the potential due to all the other molecules. The integration over the momentum space is equivalent to that for the ideal gas, see Eqn. (205). The integration over the volume splits into two parts

$$\int d\mathbf{r} e^{-\beta U(\mathbf{r})} = \int_{V_{\text{excl}}} d\mathbf{r} e^{-\infty} + \int_{V - V_{\text{excl}}} d\mathbf{r} e^{-\beta N \overline{u}} = (V - V_{\text{excl}}) e^{-\beta N \overline{u}} \tag{320}$$

Following Eqn. (118), we obtain for the partition function

$$Z_1 = n_Q (V - V_{\text{excl}}) e^{-\frac{\beta N \overline{U}}{V}}. \tag{321}$$

The free energy is then

$$F = -k_B T N \ln Z_1 = -N k_B T \left[\ln n_Q + \ln(V - V_{\text{excl}}) - \frac{N\overline{U}}{k_B T V} \right]. \tag{322}$$

The pressure is then obtained from

$$P = -\left(\frac{\partial F}{\partial V} \right)_T = N k_B T \left[\frac{1}{V - V_{\text{excl}}} + \frac{N\overline{U}}{k_B T V^2} \right]. \tag{323}$$

Rewriting gives

$$(P - \frac{\overline{U} N^2}{V^2})(V - V_{\text{excl}}) = N k_B T \tag{324}$$

The conventional Van der Waals equation is usually written in moles giving $N k_B = nR$ with $n = N/N_A$ is the number of moles of gas in the system with N_A is Avogadro's number and $R = k_B N_A$ is the ideal gas constant. Similarly, we can introduce a constant $a = -\overline{U} N_A^2$ giving $-\overline{U} N^2/V^2 = an^2/V^2$. The excluded volume $V_{\text{excl}} \sim N$ should be proportional to the number of molecules in the gas. After introducing, the constant $b = V_{\text{excl}}/n$, we arrive at the usual form of the Van der Waals equation

$$(P + \frac{an^2}{V^2})(V - bn) = nRT. \tag{325}$$

This should be compared with the ideal gas law

$$PV = nRT. \tag{326}$$

B. Critical points

Stanley Chapter 5
The pressure in Eqn. (323) in conventional notation is given by

$$P = \frac{nRT}{V - nb} - \frac{an^2}{V^2}. \tag{327}$$

For convenience, we would like to remove the n's from the equation. This can be done by considering 1 mole or let $V \to V/n$, i.e. the volume per mole. Note that for $V \gg nb$ and large T the first term dominates and we reduce to the ideal gas law. Note that the second term, resulting from the attractive interactions between the molecules effectively lowers the pressure. Although, at first sight, the Van der Waals equation appears to giving mere corrections to the ideal gas law, it displays some fundamentally different physics. Note that the pressure diverges when $V \to nb$, i.e. it becomes very difficult to compress any further. This is more characteristic of a liquid than of a gas. Unfortunately, the Van der Waals equation of state has a major deficiency. We can write the volume as

$$V^3 - (b + \frac{RT}{P})V^2 + \frac{a}{P}V - \frac{ab}{P} = 0. \tag{328}$$

This equation has three solutions, which is physically somewhat confusing. In addition, there is a region where the isotherms (V as function of p at constant T) has a positive $\partial V/\partial p$, this would correspond to a negative isothermal compressibility $K_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$. This is physically impossible since it would imply that the volume becomes larger when increasing the pressure.

??????????????????????????????????????Maxwell construction????

We are interested in obtaining the critical point. This is the point where $V_1 = V_2$ and we do not have a coexistence of phases anymore. At this point the $P - V$ isotherm has an inflection point, therefore

$$\left(\frac{\partial P}{\partial V}\right)_T = 0 \quad \text{and} \quad \left(\frac{\partial^2 P}{\partial V^2}\right)_T = 0. \tag{329}$$

We can find the critical point by solving the above equations. However, an alternative way is to state

$$(V - V_c)^3 = V^3 - 3V^2 V_c + 3V V_c^2 - V_c^3 = 0. \tag{330}$$

However, we also need to satisfy

$$V_c^3 - (b + \frac{RT_c}{P_c})V_c^2 + \frac{a}{P_c}V - \frac{ab}{P_c} = 0. \tag{331}$$

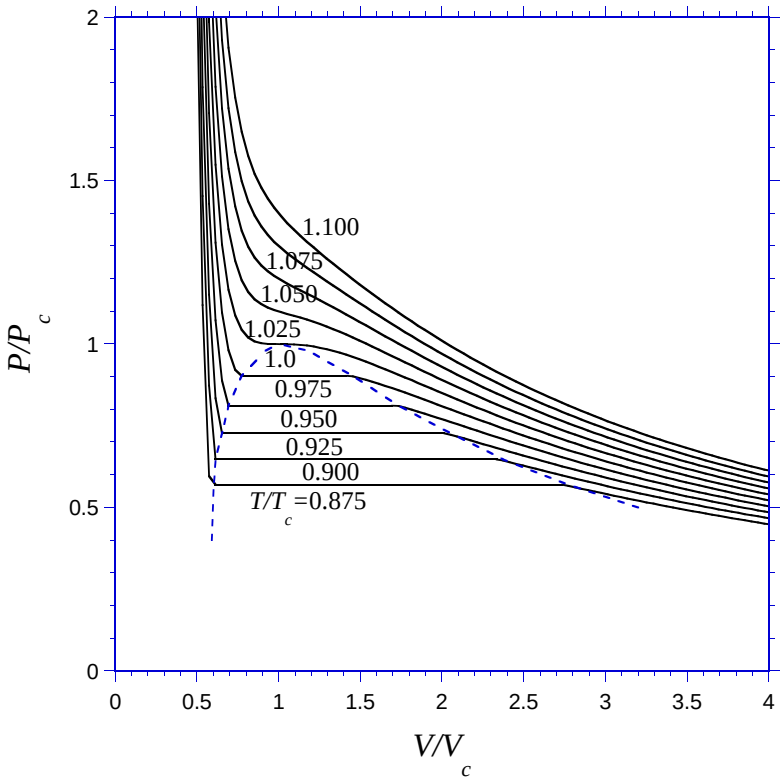

FIG. 8: A $P - V$ (pressure versus volume)

This leads to three equations

$$-3V_c = -(b + \frac{RT_c}{P_c}) \tag{332}$$

$$3V_c^2 = \frac{a}{P_c} \tag{333}$$

$$-V_c^3 = -\frac{ab}{P_c}. \tag{334}$$

Dividing the third by the second two gives

$$-\frac{V_c}{3} = -b \quad \Rightarrow \quad V_c = 3b. \tag{335}$$

Inserting gives

$$P_c = \frac{a}{27b^2}. \tag{336}$$

Inserting this into the first equation gives

$$-3(3b) = -b - \frac{27b^2 RT_c}{a} \quad \Rightarrow \quad RT_c = \frac{8a}{27b}. \tag{337}$$

Note that we can define the number

$$Z_c = \frac{P_c V_c}{RT_c} = \frac{\frac{a}{27b^2}3b}{R\frac{8a}{27b}} = \frac{3}{8}. \tag{338}$$

For ideal gases (that do no have a critical point),

$$Z = \frac{PV}{RT} = 1 \tag{339}$$

It would be nice to see if we can express the equation of state in these three parameters putting it in a universal form. This can be done by introducing the reduced variables

$$\tilde{P} = \frac{P}{P_c} \quad \text{and} \quad \tilde{V} = \frac{V}{V_c} \quad \text{and} \quad \tilde{T} = \frac{T}{T_c} \tag{340}$$

Dividing the Van der Waals Equation by $P_c V_c = \frac{3}{8}RT_c$ gives for $n = 1$ mole

$$(\tilde{P} + \frac{a}{V^2}\frac{27b^2}{a})(\tilde{V} - \frac{b}{3b}) = \frac{8}{3}\tilde{T}, \tag{341}$$

or

$$(\tilde{P} + \frac{3}{\tilde{V}^2})(3\tilde{V} - 1) = 8\tilde{T}. \tag{342}$$

Note that the Van der Waals equation is now expressed solely in the critical values of the gas. We would now like to express the behavior of the gas close to that critical point. We introduce the variables that approach zero when approaching the critical point

$$p = \tilde{P} - 1 = \frac{P}{P_c} - 1 \quad \text{and} \quad v = \tilde{V} - 1 = \frac{V}{V_c} - 1 \quad \text{and} \quad \tau = \tilde{T} - 1 = \frac{T}{T_c} - 1. \tag{343}$$

We want to rewrite Eqn. (342) in the new variables

$$(\tilde{P}\tilde{V}^2 + 3)(3\tilde{V} - 1) = 8\tilde{T}\tilde{V}^2 \tag{344}$$

$$\tilde{P}(3\tilde{V}^3 - \tilde{V}^2) + 9\tilde{V} - 3 - 8\tilde{V}^2 = 8(\tilde{T} - 1)\tilde{V}^2 \tag{345}$$

$$p(3\tilde{V}^3 - \tilde{V}^2) + 3\tilde{V}^3 - 9\tilde{V}^2 + 9\tilde{V} - 3 = 8\tau\tilde{V}^2 \tag{346}$$

$$p[3(v+1)^3 - (v+1)^2] + 3v^3 = 8\tau(v+1)^2 \tag{347}$$

yielding

$$p(3v^3 + 8v^2 + 7v + 2) + 3v^3 \;=\; 8\tau(v^2 + 2v + 1). \tag{348}$$

First, let us consider the behavior along the critical isotherm $\tau = 0$. Since $v \ll 1$, we obtain

$$2p + 3v^3 \;=\; 0 \quad \Rightarrow \quad p = -\frac{3}{2}v^3. \tag{349}$$

This is the degree of flatness of the critical isotherm at the critical point. Next, we want to examine the dependence of the volume as it approaches the critical point from below. Let us therefore express Eqn. (V B) as a polynomial in v

$$3(1+p)v^3 + 8(p-\tau)v^2 + (7p - 16\tau)v + 2(p - 4\tau) = 0. \tag{350}$$

We can neglect the first p since $p \ll 1$. This gives three solutions v_1, v_2, and v_3 comparable to what we know from the Maxwell construction. Therefore one is much smaller than the other two $v_2 \ll v_{1,3}$ and corresponds to the solution in the coexistence region. We are more interested in the other two reduced volumes that should be comparable in magnitude $|v_1| \cong |v_2|$. The means that the equation should be of the form $v(v^2 - a) = v(v - a)(v + a) = 0$, where we would like to determine a. To get the equation in this form implies that

$$p \cong 4\tau \tag{351}$$

Using this gives

$$3v^3 + 24\tau v^2 + 12\tau v = 0 \quad \Rightarrow \quad v(v^2 + 8\tau v + 4\tau) \cong 0 \quad \Rightarrow \quad v^2 + 4\tau(1 + 2v) \cong v^2 + 4\tau \cong 0, \tag{352}$$

yielding

$$v \cong \pm 2|\tau|^{\frac{1}{2}}, \tag{353}$$

where the negative value corresponds to the liquid and the positive one to the gas. We can also rewrite this in terms of the density

$$\frac{\rho}{\rho_c} - 1 = \frac{V_c}{V} - 1 = \frac{V_c}{V_c(1+v)} - 1 = \frac{1}{1+v} - 1 \cong 1 - v - 1 = -v, \tag{354}$$

which is just the opposite of the reduced volume.

$$\frac{\rho}{\rho_c} - 1 \cong \mp 2|\tau|^{\frac{1}{2}}, \tag{355}$$

for gas and liquid near the critical point, respectively.

We would also like to consider the compressibility of the system. In reduced variables, this is given by

$$\kappa = -\left(\frac{\partial v}{\partial p}\right)_\tau. \tag{356}$$

Differentiating Eqn. () with respect to p gives

$$(3v^3 + 8v^2 + 7v + 2) + p(9v^2 + 16v + 7)\frac{\partial v}{\partial p} + 9v^2\frac{\partial v}{\partial p} \;=\; 8\tau(2v + 2)\frac{\partial v}{\partial p}. \tag{357}$$

Keeping only the dominant terms gives

$$2 + 7p\frac{\partial v}{\partial p} + 9v^2\frac{\partial v}{\partial p} \;=\; 16\tau\frac{\partial v}{\partial p} \tag{358}$$

which gives

$$-\left(\frac{\partial v}{\partial p}\right)_\tau = \frac{2}{7p + 9v^2 - 16\tau}. \tag{359}$$

For $\tau > 0$, we approach the critical point in several ways. Keeping the volume constant (isochore) at the critical volume, $v = 0$, gives, using Eqn. (351)

$$-\left(\frac{\partial v}{\partial p}\right)_{\tau\to 0^+} = \frac{2}{28\tau - 16\tau} = \frac{1}{6\tau}. \tag{360}$$

For $\tau < 0$, we approach the critical point along the co-existence curve for which $v^2 = -4\tau$, giving

$$-\left(\frac{\partial v}{\partial p}\right)_{\tau\to 0^+} = \frac{2}{28\tau - 36\tau - 16\tau} = -\frac{2}{24\tau} = \frac{1}{12|\tau|}. \tag{361}$$

Since the calculation is somewhat lengthy, we only give the results for the specific heat

$$\Delta C_V = C_V - C_V^0 = \begin{cases} \frac{9}{2}Nk_B[1 - \mathcal{O}(\tau)] & T < T_c \\ 0 & T > T_c \end{cases}, \tag{362}$$

where $C_V^0 = \frac{3}{2}Nk_B$ is the constant volume specific heat predicted for the noninteracting ideal gas. Therefore for the dependence $C_V \sim (1 - T/T_c)^{\alpha'}$ for $T < T_c$ and $C_V \sim (T/T_c - 1)^{\alpha}$ for $T > T_c$, we find $\alpha = \alpha' = 0$ (actually for $T < T_c$ this expression is not well defined).

C. The Ising model: mean field

The Ising model is given by the Hamiltonian

$$H = -J\sum_{\langle ij\rangle}\sigma_i\sigma_j - \mu_B B\sum_{i=1}^{N}\sigma_i, \tag{363}$$

where the sum over $\langle ij\rangle$ goes over all the nearest neighbor spins. This is our two-level system, but now with interacting spins. Note that for $J > 0$, the spins prefer to be parallel so we can expect ferromagnetism.
See Baxter Chapter 3

Let us consider the Hamiltonian in Eqn. (363) in a slightly different fashion

$$H = \sum_{i=1}^{N}\left\{-J\sum_{\text{n.n. } j}\sigma_j - \mu_B B\right\}\sigma_i, \tag{364}$$

where the summation goes over the nearest neighbor (n.n.) j's. This implies that we can also view the Hamiltonian as an effective field working on the spins

$$\mu_B B_{\text{eff}} = \mu_B B + J\sum_{\text{n.n. } j}\sigma_j. \tag{365}$$

Unfortunately this is still not very practical since we have no idea what the spins σ_j are. The mean-field approximation starts from the idea that all spins are equivalent and we should obtain a good idea of the behavior of a spin at site j by taking an average ("mean") behavior of all the spins. This gives an effective field

$$\mu_B B_{\text{eff}} = \mu_B B + J\frac{q}{N-1}\sum_{j\neq i}\sigma_j, \tag{366}$$

where q is the number of nearest neighbors. In that limit, we can rewrite the Hamiltonian as

$$H = -J\frac{q}{N-1}\sum_{(ij)}\sigma_i\sigma_j - \mu_B B\sum_{i=1}^{N}\sigma_i, \tag{367}$$

where the summation (ij) goes over all the $\frac{1}{2}N(N-1)$ possible pairs. Let us rewrite the interaction term

$$\sum_{(ij)}\sigma_i\sigma_j = \frac{1}{2}\sum_{i,j,i\neq j}\sigma_i\sigma_j = \frac{1}{2}(\sum_{i,j}\sigma_i\sigma_j - \sum_i\sigma_i\sigma_i) = \frac{1}{2}(\sum_i\sigma_i\sum_j\sigma_j - N) = \frac{1}{2}(\mathcal{M}^2 - N), \tag{368}$$

where the $\frac{1}{2}$ accounts for the double counting when replacing the sum over the pairs by the sum over the spins. We also used $\sigma_i^2 = 1$ and $\mathcal{M} = \sum_i \sigma_i$. Note that $M = \mu_B \mathcal{M}$. The total energy can therefore be written in terms of the magnetization

$$E(\mathcal{M}) = -J\frac{q}{2(N-1)}(\mathcal{M}^2 - N) - \mu_B B\mathcal{M}. \tag{369}$$

Let us express the magnetization in the number of spins. If we have r down spins, and therefore $N - r$ up spins, the magnetization is given by

$$\mathcal{M} = (N-r) \times 1 + r \times (-1) = N - 2r. \tag{370}$$

We can then express the partition function as a sum over the number of down spins (or effectively the magnetization) as opposed to a sum over all the spin configurations:

$$Z = \sum_{r=0}^{N} c_r, \tag{371}$$

where

$$c_r = \frac{N!}{r!(N-r)!} \exp\left\{\beta J \frac{q}{2(N-1)}[(N-2r)^2 - N] + \beta\mu_B B(N-2r)\right\}, \tag{372}$$

where the factor $\frac{N!}{r!(N-r)!}$ account for the number of possible states with r down spins.
Typical behaviors of c_r are given in Fig. 9. The values of β would correspond to very high temperatures. However, we set J equal to unity. This is unphysically large, since J should be of the order of $k_B T_C$ where T_C is the transition temperature. Since $k_B T_{\text{room}} \cong 25$ meV, J should be of the order of magnitude since critical temperature range from 0 K up to several hundred Kelvin. Another way to look at it is that the temperature scale is given in terms of J. This is often done in theoretical papers. $\beta = 0.2$ then corresponds to a temperature of $5J$, which is far above the the critical temperature (independent of exact value of the critical temperature, be it 2 or 600 K). Note that for small $\beta = 1/k_B T$, i.e. high temperatures, the exponent approaches unity and c_r is determined by the probability factor $\frac{N!}{r!(N-r)!}$. For lower temperatures or larger values of β, the energy $E(\mathcal{M})$ starts to dominate and parallel spins are preferable. We observe two maxima since the spins can be either all up or all down. Note, that one maximum is higher due to the presence of a magnetic field B. The calculations are here done for only 80 spins. The maxima would be significantly sharper for a more reasonable number of spins (or the order of 10^{23}). For $\beta = 0.5$, we already approach full magnetization.

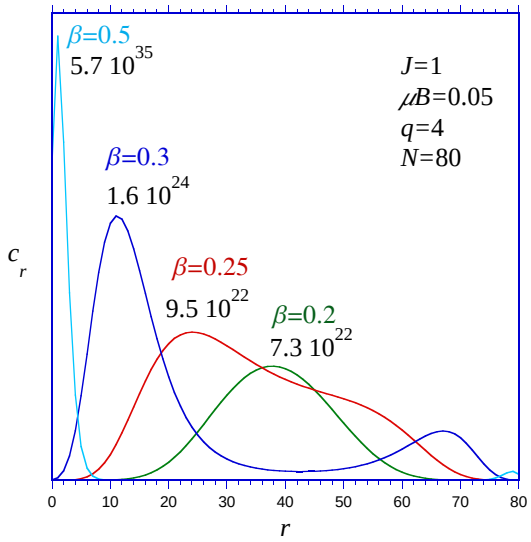

FIG. 9: Typical curves for the behavior of c_r. Note that the plots are not to scale and that the maximum values is given in the figure. The parameters used are in the figure.

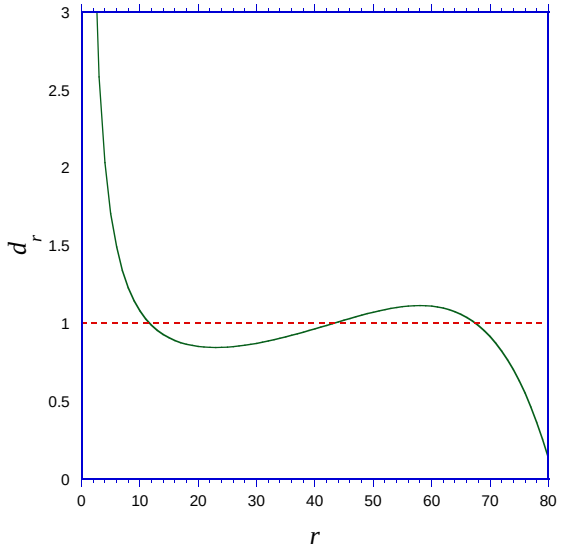

FIG. 10: The value of d_r for $\beta = 0.3$ using the parameters given in the Figure for c_r. Note that there are three values of r that give $d_r = 1$.

Let us now consider the ratios $d_r = c_{r+1}/c_r$. Since there is a maximum at some point this ratio much change from greater to one to less than one for a particular r_m. For very large number of particles, the function becomes continuous and the maximum is given by $d_{r_m} = 1$. The ratio is given by

$$\begin{aligned} d_r &= \frac{c_{r+1}}{c_r} = \frac{r!(N-r)!}{(r+1)!(N-r-1)!} e^{\beta J \frac{q}{2(N-1)}[(N-2r-2)^2-N]+\beta\mu_B B(N-2r-2)-\beta J \frac{q}{2(N-1)}[(N-2r)^2-N]+\beta\mu_B B(N-2r)} \\ &= \frac{N-r}{r+1} \exp\left\{-\beta J \frac{2q}{N-1}[N-2r+1] - 2\beta\mu_B B\right\} \end{aligned} \tag{373}$$

This function is plotted in Fig. 10. Note that the extrema in c_r are given by the values of r for which $d_r = 1$. The magnetization is given by $\mathcal{M} = N - 2r$. Therefore, the magnetization per site is $m = \mathcal{M}/N = 1 - 2r_m/N$. For large

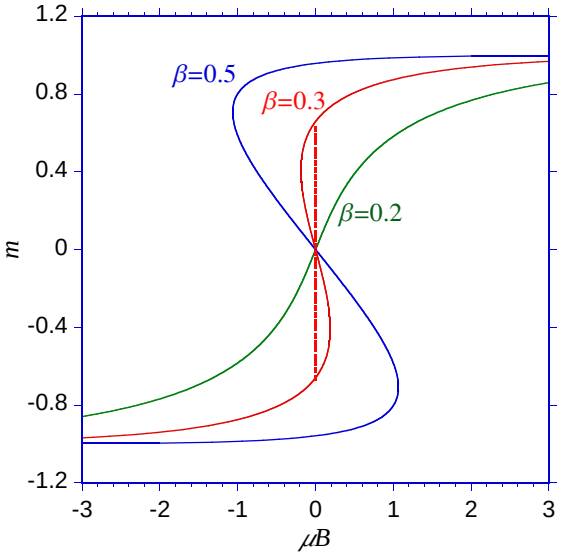

FIG. 11: The magnetization per site m as a function of the magnetic field $\mu_B B$ for the parameters given in Fig. 9.

N, $N-1 \cong N$ and we can write the previous equation as

$$d_{r_m} = 1 = \frac{1+m}{1-m} e^{-2\beta q J m - 2\beta \mu_B B}. \tag{374}$$

Let us denote $a = 2\beta q J m + 2\beta \mu_B B$. Solving for m gives

$$1 - m = e^{-a} + m e^{-a} \quad \Rightarrow \quad m(1+e^{-a}) = 1 - e^{-a} \quad \Rightarrow \quad m = \frac{1-e^{-a}}{1+e^{-a}} = \frac{e^{a/2} - e^{-a/2}}{e^{a/2} + e^{-a/2}} = \tanh \frac{a}{2}, \tag{375}$$

or

$$m = \tanh(\beta q J m + \beta \mu_B B) = \tanh \frac{q J m + \mu_B B}{k_B T}. \tag{376}$$

This result was first obtained by Bragg and Williams in 1934. This is the mean-field Hamiltonian. In some sense, we are solving the problem in infinite dimensions since all spins are effectively surrounded by an infinite number of nearest neighbors with the same interaction strength. However, the interaction strength depends on the number of particles, which is rather unphysical. Surprisingly, the simple model often gives a reasonable description of the thermodynamic properties.

We can also rewrite this as

$$\mu_B B = -q J m + k_B T \text{arctanh } m = -q J m + \frac{1}{\beta} \text{arctanh } m. \tag{377}$$

This allows us to plot m as a function of $\mu_B B$, see Fig. 11 for several values of β. For large temperatures ($\beta = 0.2$), the magnetization is zero. For lower temperatures, we see that the magnetization can have three possible values. However, the three possible values correspond to the extrema in c_r, see Fig. 9. From Fig. 10, we also see that there are three r values for which $d_r = 1$. However, the one at $M = 0$ is in fact a minimum, so the only possible maxima are the values at a finite magnetization. The direction of the magnetization depends on the direction of the magnetic field. As we can see clearly from Fig. 9, there is only one maximum with the spins parallel to the magnetization (Note that the curve will reverse when the direction of the magnetization is changed). Therefore, for $0 < \mu_B B \ll 1$, the magnetization m is positive, but for small $\mu_B B > 0$ the magnetization m is negative. There is therefore a jump in the magnetization at $\mu_B B = 0$. This discrete jumbp at $B = 0$ is indicated by the dashed line in Fig. 11. The behavior given Eqn. (377) corresponds to spurious extrema and has no physical meaning (this construction is similar to the Maxwell construction for the Van der Waals gas). In real systems, the magnetization is often pinned in a certain direction due to the presence of anisotropies or impurities. In that case, when going from a positive to a negative magnetic field, the magnetization will remain positive. On the other hand, it requires a finite positive magnetic field to flip the magnetization from negative to positive.

Spontaneous magnetization.– We have seen that we can obtain a finite magnetization at a certain temperature even in the absence of a magnetic field. For magnetism, the magnetization is the order parameter that indicates a particular type of (spontaneous) long range order. For $B = 0$,

$$m = \tanh \frac{q J m}{k_B T}. \tag{378}$$

At high temperatures, the magnetization is zero. Therefore there must be a temperature, known as the critical temperature T_c, where spontaneous magnetization occurs. Close to T_c, the magnetization m must be small and we can write

$$m = \frac{q J m}{k_B T}, \tag{379}$$

using $\tanh x \cong x$ for $x \to 0$. When $T \cong T_c$ this can only be correct if

$$T_c = \frac{q J}{k_B}. \tag{380}$$

We can also describe the behavior of the magnetization as a function of temperature using

$$m = \tanh \frac{T_c m}{T} \quad \Rightarrow \quad m = \frac{T}{T_c} \text{arctanh } m. \tag{381}$$

Using arctanh $m \cong m + \frac{1}{3}m^3$, we obtain

$$m = \frac{T}{T_c}(m + \frac{1}{3}m^3) \quad \Rightarrow \quad m(\frac{T}{T_c} - 1 + \frac{1}{3}m^2) = 0. \tag{382}$$

This gives as possible solutions $m = 0$, which we know to be a minimum in the probability, and

$$m = \sqrt{3(1 - \frac{T}{T_c})}(1 + \mathcal{O}(T)), \tag{383}$$

where the $\mathcal{O}(t)$ indicates higher-order term that we did not include.

Often one is interested in the critical behavior of certain quantities when approaching the critical temperature. Some quantities, such as the order parameter approach zero when $T \uparrow T_c$. Other quantities however can diverge. Different models can give different behaviors. The behavior of the magnetization close to T_c is defined as

$$m \sim (1 - \frac{T}{T_c})^{\beta}, \tag{384}$$

with β the standard notation for the critical exponent for the order parameter.

Susceptibility. – Often we are aloso interested in the change in magnetic field as a function of temperature. The susceptibility per site is defined as

$$\chi = \left(\frac{\partial m}{\partial H}\right)_T = \frac{1}{\mu_B}\left(\frac{\partial m}{\partial B}\right)_T. \tag{385}$$

Differentiating Eqn. (377) with respect to B gives

$$\mu_B = -qJ\frac{\partial m}{\partial B} + k_B T \frac{1}{1 - m^2}\frac{\partial m}{\partial B} = \frac{\partial m}{\partial B}\frac{-qJ(1 - m^2) + k_B T}{1 - m^2} \tag{386}$$

using the derivative $(\text{arctanh } m)' = (1 - m^2)^{-1}$. This gives for the susceptibility

$$\chi = \frac{1}{\mu_B}\frac{\partial m}{\partial B} = \frac{1 - m^2}{qJ(\frac{T}{T_c} - 1 + m^2)}. \tag{387}$$

Let us first consider the situation $T > T_c$, for which $m = 0$, giving

$$\chi = \frac{1}{qJ(\frac{T}{T_c} - 1)}. \tag{388}$$

For $T_c < T$, we derived above that $m^2 = 3(1 - \frac{T}{T_c})$. Inserting this gives

$$\chi = \frac{1 - 3(1 - \frac{T}{T_c})}{qJ(\frac{T}{T_c} - 1 + 3(1 - \frac{T}{T_c}))} = \frac{-2 + 3\frac{T}{T_c}}{2qJ(1 - \frac{T}{T_c})} \cong \frac{1}{2qJ(1 - \frac{T}{T_c})}, \tag{389}$$

for $T/T_c \cong 1$. The critical exponents for the susceptibility are defined as

$$\chi \sim \frac{1}{(\frac{T}{T_c} - 1)^{\gamma}} \qquad T > T_c \tag{390}$$

$$\chi \sim \frac{1}{(1 - \frac{T}{T_c})^{\gamma'}} \qquad T < T_c. \tag{391}$$

In the mean-field approximation, we therefore find that $\gamma = \gamma' = 1$. We find that the susceptibility diverges as we approach the critical temperature. This can be understood as follows. Close to T_c the magnetization is small. Therefore, a magnetic field can cause a large change in the magnetization. However, far above T_c, temperature fluctuations dominate and the response to a magnetic field decreases. Far below T_c, the system is fully magnetized and a magnetic field will have only a small effect on the magnetization.

For the specific heat, we first need to determine the free energy. The partition function is given in Eqn. (371) as a summation over the different magnetizations r. When N is large, we can assume that the partition function is dominated by the values close to r_0. This gives for the free energy per site

$$F = -\frac{1}{N}k_BT\ln Z \cong -\frac{1}{N}k_BT\ln c_{r_0}. \tag{392}$$

The calculation is somewhat lengthy and will be skipped here, but gives for the free energy per site

$$F = -\frac{1}{2}k_BT\ln\frac{4}{1-m^2} + \frac{1}{2}qJm^2. \tag{393}$$

For $T > T_c$, the magnetization is zero and the free energy becomes

$$F = -k_BT\ln 2, \tag{394}$$

where the $\ln 2$ corresponds to the two spin directions at a particular site. The specific heat is therefore zero. Since the system is completely disordered in the absence of a magnetic field, the internal energy does not change as a function of temperature.

For $T < T_c$, we have a finite magnetization. Close to T_c, it is small and we have

$$\begin{aligned} F &= -k_BT\ln 2 + \frac{1}{2}k_BT\ln(1-m^2) + \frac{1}{2}qJm^2 \cong -k_BT\ln 2 - \frac{1}{2}k_BTm^2 + \frac{1}{2}qJm^2 - \frac{1}{4}k_BTm^4 &(395)\\ &= -k_BT\ln 2 - \frac{1}{2}k_BTm^2(1-\frac{qJ}{k_BT}) - \frac{1}{4}k_BTm^4, &(396)\end{aligned}$$

using $\ln(1-m^2) = (-m^2) - \frac{1}{2}(-m^2)^2 + \cdots$. Since $m^2 = 3(1-T/T_c)$, we obtain

$$\begin{aligned} F &= -k_BT\ln 2 - \frac{1}{2}k_BT3(1-\frac{T}{T_c})(1-\frac{T_c}{T}) - \frac{9}{4}k_BT\left(1-\frac{T}{T_c}\right)^2 &(397)\\ &= -k_BT\ln 2 + \frac{3}{2}k_BT\frac{T}{T_c}\left(1-\frac{T}{T_c}\right)^2 - \frac{9}{4}k_BT\left(1-\frac{T}{T_c}\right)^2 &(398)\\ &= -k_BT\ln 2 - \frac{3}{4}k_BT\left(1-\frac{T}{T_c}\right)^2 &(399)\end{aligned}$$

In the neighborhood of T_c, we can write approximately

$$C_V \cong -T_c\frac{\partial^2}{\partial T^2}\left[-\frac{3}{4}k_BT_c\left(1-\frac{T}{T_c}\right)^2\right] = T_c\frac{\partial}{\partial T}\left[\frac{3}{2}k_BT_c\left(1-\frac{T}{T_c}\right)\left(-\frac{1}{T_c}\right)\right] = \frac{3}{2}k_B. \tag{400}$$

We see that that specific heat is again constant but nonzero. We therefore find for the critical exponent $\alpha = \alpha' = 0$.

D. Relations between liquid-gas phase transition and magnetism

We started out by discussing phase transitions from

E. Ising model: exact solution in one Dimension

See Pathria Chapter 12 and Baxter Chapter 2

There are relatively few systems that can be solved exactly. One of these is the Ising model and up till now only in one and two dimensions. Unfortunately, the solution for two dimensions is already complicated so let us just stick to that for one dimension since that already gives some surprising results. We will treat this Hamiltonian in one dimension and use periodic boundary conditions. Effectively, this means that the spins are lying on a ring with the N'th spin attached to the first one. However, "periodic boundary conditions" is a mathematical device and certainly more difficult to imagine for higher dimensions. The assumption is that for a very large system the boundaries are

not really relevant. Let us first rewrite the Hamiltonian from Eqn. (363) in a somewhat more symmetric convenient form

$$H = -J\sum_i \sigma_i\sigma_{i+1} - \frac{1}{2}\mu_B B\sum_{i=1}^{N}(\sigma_i + \sigma_{i+1}). \tag{401}$$

Remember, that for the two-level system, we were able to rewrite $Z = Z_1^N$. This was relatively obvious since each spin was acting independently. Here we would like to do the same thing but unfortunately neighboring spins are no longer independent. A solution for this was found by Kramers and Wannier who used a matrix method. Formally we can write

$$Z_N = \sum_{\sigma_i=\pm 1} \langle\sigma_1|U_\sigma|\sigma_2\rangle\langle\sigma_2|U_\sigma|\sigma_3\rangle\cdots\langle\sigma_N|U_\sigma|\sigma_1\rangle. \tag{402}$$

Each of the terms is given by

$$\langle\sigma_i|U_\sigma|\sigma_{i+1}\rangle = \exp\left\{\beta[J\sigma_i\sigma_{i+1} + \frac{1}{2}\mu_B B(\sigma_i + \sigma_{i+1})]\right\}. \tag{403}$$

This can be achieved if we define the matrix U

$$U_\sigma = \begin{pmatrix} e^{\beta(J+\mu_B B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu_B B)} \end{pmatrix} \tag{404}$$

and we use for the different $|\sigma\rangle$

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{405}$$

Since we have written things in terms of matrices, we can use some nice properties of matrix algebra. First, we note that the spin states form a complete set

$$\sum_i |\sigma_i\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(0,1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \tag{406}$$

and is therefore equal to a unity matrix. The partition function therefore greatly simplifies

$$\begin{aligned} Z_N &= \sum_{\sigma_1}\langle\sigma_1|U_\sigma\sum_{\sigma_2}|\sigma_2\rangle\langle\sigma_2|U_\sigma\sum_{\sigma_3}|\sigma_3\rangle\cdots\langle\sigma_N|U_\sigma|\sigma_1\rangle & (407)\\ &= \sum_{\sigma_1}\langle\sigma_1|U_\sigma I U_\sigma I\cdots I U_\sigma|\sigma_1\rangle & (408)\\ &= \sum_{\sigma_1}\langle\sigma_1|U_\sigma^N|\sigma_1\rangle. & (409)\end{aligned}$$

The multiplication is somewhat cumbersome, and much more easily done if the matrix is diagonal. The matrix is easily diagonalized

$$\begin{vmatrix} e^{\beta(J+\mu_B B)} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu_B B)} - \lambda \end{vmatrix} = (e^{\beta(J+\mu_B B)} - \lambda)(e^{\beta(J-\mu_B B)} - \lambda) - e^{-2\beta J} = 0 \tag{410}$$

or

$$\begin{aligned} \lambda^2 - \lambda(e^{\beta(J+\mu_B B)} + e^{\beta(J-\mu_B B)}) + e^{2\beta J} - e^{-2\beta J} &= & (411)\\ \lambda^2 - 2\lambda e^{\beta J}\cosh(\beta\mu_B B) + 2\sinh(2\beta J) &= 0. & (412)\end{aligned}$$

This gives the eigenvalues

$$\lambda_\pm = e^{\beta J}\cosh(\beta\mu_B B) \pm \frac{1}{2}\sqrt{4e^{2\beta J}\cosh^2(\beta\mu_B B) - 8\sinh(2\beta J)}. \tag{413}$$

The square root can be rewritten as

$$\sqrt{e^{2\beta J}\cosh^2(\beta\mu_B B) - 2\sinh(2\beta J)} = \sqrt{e^{2\beta J}\cosh^2(\beta\mu_B B) - (e^{2\beta J} - e^{-2\beta J})} = \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}}, \quad (414)$$

using $\cosh^2 x - \sinh^2 x = 1$, giving

$$\lambda_\pm = e^{\beta J}\cosh(\beta\mu_B B) \pm \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}} \quad (415)$$

The new matrixs is diagonal

$$U_\lambda = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}. \quad (416)$$

The partition function now becomes

$$Z_N = \sum_{\lambda_1}\langle\lambda_1|U_\lambda^N|\lambda_1\rangle = \sum_{\lambda_1}\langle\lambda_1|\begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}^N|\lambda_1\rangle = \sum_{\lambda_1}\langle\lambda_1|\begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix}|\lambda_1\rangle = \lambda_+^N + \lambda_-^N. \quad (417)$$

Let us first consider the situation that $J = 0$. In this limit, we find

$$\lambda_\pm = \cosh(\beta\mu_B B) \pm \sqrt{\sinh^2(\beta\mu_B B) + 1} = \cosh(\beta\mu_B B) \pm \cosh(\beta\mu_B B) = \begin{cases} 2\cosh(\beta\mu_B B) \\ 0 \end{cases} \quad (418)$$

This gives the partition function $Z_N = [2\cosh(\beta\mu_B B)]^N$, which reproduces the results in Eqns. (82) and (83).

Since $\lambda_- < \lambda_+$, we generally have $(\lambda_-/\lambda_+)^N \to 0$ as $N \to \infty$. The properties are therefore determined by λ_+. The free energy is then

$$F = -k_B T \ln Z_N \cong -k_B T N \ln\lambda_+ = -k_B T N \ln\left[e^{\beta J}\cosh(\beta\mu_B B) + \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}}\right] \quad (419)$$

$$= -NJ - Nk_B T \ln\left[\cosh(\beta\mu_B B) + \sqrt{\sinh^2(\beta\mu_B B) + e^{-4\beta J}}\right]. \quad (420)$$

In particular, we can derive the magnetization per site by using Eqn. (103),

$$m = \frac{M}{\mu_B N} = -\frac{1}{N}\left(\frac{\partial F}{\partial B}\right)_T \quad (421)$$

$$= k_B T\frac{\partial}{\partial B}\ln\left[e^{\beta J}\cosh(\beta\mu_B B) + \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}}\right]$$

$$= k_B T\frac{1}{e^{\beta J}\cosh(\beta\mu_B B) + \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}}}\left[e^{\beta J}\sinh(\beta\mu_B B) + \frac{1}{2}\frac{e^{2\beta J}2\sinh(\beta\mu_B B)\cosh(\beta\mu_B B)}{\sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}}}\right]\beta\mu_B,$$

which gives

$$M = \frac{\mu_B N e^{\beta J}\sinh(\beta\mu_B B)}{\sqrt{e^{2\beta J}\sinh^2(\beta\mu_B B) + e^{-2\beta J}}}. \quad (422)$$

This gives $M = 0$ for $B = 0$ and hence there is no spontaneous magnetization as we have found using the mean-field approximation. Note however that at $T = 0$, an infinitesimally small magnetic field gives a magnetization of $\mu_B N$. Therefore, the critical temperature is $T = 0$, again in contrast with the mean-field approximation which gives $T_c = \frac{\mu_B J}{k_B}$.

correlation functions.–

We can also use this approach to calculatate correlation functions, such as, for example,

$$\langle\sigma_1\sigma_3\rangle = \frac{1}{Z_N}\sum_{\sigma_i=\pm1}\sigma_1\langle\sigma_1|U_\sigma|\sigma_2\rangle\langle\sigma_2|U_\sigma|\sigma_3\rangle\sigma_3\langle\sigma_3|U_\sigma|\sigma_4\rangle\cdots\langle\sigma_N|U_\sigma|\sigma_1\rangle. \quad (423)$$

This can also be written as a matrix with matrix elements

$$S_{\sigma,\sigma'} = \sigma\delta_{\sigma,\sigma'}, \tag{424}$$

or

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{425}$$

Again removing all the spin by completeness gives

$$\langle\sigma_1\sigma_3\rangle = \frac{1}{Z_N}\sum_{\sigma_1}\langle\sigma_1|SU_\sigma U_\sigma SU\cdots U_\sigma|\sigma_1\rangle. \tag{426}$$

We can easily generalize these concepts to give

$$\langle\sigma_i\sigma_j\rangle = \frac{1}{Z_N}\sum_{\sigma_1}\langle\sigma_1|SU_\sigma^{j-i}SU_\sigma^{N-j+i}|\sigma_1\rangle \tag{427}$$

$$m = \langle\sigma_i\rangle = \frac{1}{Z_N}\sum_{\sigma_1}\langle\sigma_1|SU_\sigma^{N}|\sigma_1\rangle. \tag{428}$$

Let us return for a second to what we did in the previous section. We made the matrix U diagonal. Effectively, we can say we performed a unitary transformation to a new basis set

$$|\lambda_+\rangle = \cos\varphi|\uparrow\rangle + \sin\varphi|\downarrow\rangle \tag{429}$$

$$|\lambda_-\rangle = -\sin\varphi|\uparrow\rangle + \cos\varphi|\downarrow\rangle. \tag{430}$$

Note that the diagonal matrix element for $|\downarrow\rangle$ is smaller than that for $|\uparrow\rangle$. The lowest eigenvalues therefore has mainly $|\downarrow\rangle$ character. The minus is needed for the lowest eigenstate since the off-diagonal matrix element is larger than zero. The angle φ is determined by

$$\tan 2\varphi = \frac{2U_\sigma(\uparrow,\downarrow)}{U_\sigma(\uparrow,\uparrow) - U_\sigma(\downarrow,\downarrow)} = \frac{2e^{-\beta J}}{e^{\beta J}(e^{\beta\mu_B B} - e^{-\beta\mu_B B})} = \frac{1}{e^{2\beta J}\sinh(\beta\mu_B B)}. \tag{431}$$

The unitary transformation from the basis set $|\lambda_\pm\rangle$ to the basis set $|\sigma\rangle$ is given by

$$A_{\sigma\lambda} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \tag{432}$$

and vice-versa

$$A_{\lambda\sigma} = A^\dagger_{\sigma\lambda} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \tag{433}$$

Note that

$$A_{\lambda\sigma}A^\dagger_{\sigma\lambda} = I. \tag{434}$$

The matrices are related to each other by

$$U_\lambda = A_{\lambda\sigma}U_\sigma A_{\sigma\lambda} \tag{435}$$

So more formally, what we did in the previous section corresponded to

$$\sum_{\sigma_i}\langle\sigma_i|A_{\sigma\lambda}A_{\lambda\sigma}U_\sigma A_{\sigma\lambda}A_{\lambda\sigma}U_\sigma A_{\sigma\lambda}\cdots A_{\lambda\sigma}U_\sigma A_{\sigma\lambda}A_{\lambda\sigma}|\sigma_i\rangle = \sum_{\lambda_i}\langle\lambda_i|U_\lambda U_\lambda\cdots U_\lambda|\lambda_i\rangle = \sum_{\lambda_i}\langle\lambda_i|U_\lambda^N|\lambda_i\rangle. \tag{436}$$

However, now we are interested in correlation function. Applying the same procedure gives

$$\langle\sigma_i\rangle = \frac{1}{Z_N}\sum_{\sigma_i}\langle\sigma_i|A_{\sigma\lambda}A_{\lambda\sigma}SA_{\sigma\lambda}A_{\lambda\sigma}U_\sigma A_{\sigma\lambda}\cdots A_{\lambda\sigma}U_\sigma A_{\sigma\lambda}A_{\lambda\sigma}|\sigma_i\rangle \tag{437}$$

$$= \sum_{\lambda_i}\langle\lambda_i|S_\lambda U_\lambda U_\lambda\cdots U_\lambda|\lambda_i\rangle, \tag{438}$$

with S_λ given by

$$S_\lambda = A_{\lambda\sigma} S A_{\sigma\lambda} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \tag{439}$$

$$= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ -\sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} \cos^2\varphi - \sin^2\varphi & -2\sin\varphi\cos\varphi \\ -2\sin\varphi\cos\varphi & \sin^2\varphi - \cos^2\varphi \end{pmatrix} \tag{440}$$

$$= \begin{pmatrix} \cos 2\varphi & -\sin 2\varphi \\ -\sin 2\varphi & -\cos 2\varphi \end{pmatrix} \tag{441}$$

$$\langle\sigma_i\rangle = \frac{1}{Z_N}\sum_{\lambda_i}\langle\lambda_i|S_\lambda U_\lambda^N|\lambda_i\rangle = \frac{1}{Z_N}\sum_{\lambda_i}\langle\lambda_i| \begin{pmatrix} \cos 2\varphi & -\sin 2\varphi \\ -\sin 2\varphi & -\cos 2\varphi \end{pmatrix} \begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} |\lambda_i\rangle \tag{442}$$

$$= \frac{1}{Z_N}\sum_{\lambda_i}\langle\lambda_i| \begin{pmatrix} \cos 2\varphi\lambda_+^N & -\sin 2\varphi\lambda_-^N \\ -\sin 2\varphi\lambda_+^N & -\cos 2\varphi\lambda_-^N \end{pmatrix} |\lambda_i\rangle = \frac{\cos 2\varphi(\lambda_+^N - \lambda_-^N)}{\lambda_+^N + \lambda_-^N}. \tag{443}$$

Since $\lambda_-/\lambda_+^N \to 0$ for very large N, we obtain

$$\langle\sigma_i\rangle = \cos 2\varphi. \tag{444}$$

Now since $\tan 2\varphi = 1/a$, with $a = e^{2\beta J}\sinh(\beta\mu_B B)$, we have

$$\cos 2\varphi = \frac{a}{\sqrt{1+a^2}} = \frac{e^{2\beta J}\sinh(\beta\mu_B B)}{\sqrt{1 + e^{4\beta J}\sinh^2(\beta\mu_B B)}} = \frac{e^{\beta J}\sinh(\beta\mu_B B)}{\sqrt{e^{-2\beta J} + e^{2\beta J}\sinh^2(\beta\mu_B B)}}, \tag{445}$$

reproducing the result in Eqn. (422). Let us now turn our attention to the other correlation functions

$$\langle\sigma_i\sigma_j\rangle = \frac{1}{Z_N}\sum_{\sigma_1}\langle\lambda_+|S_\lambda U_\lambda^{j-i} S_\lambda U_\lambda^{N-j+i}|\lambda_-\rangle \tag{446}$$

$$= \frac{1}{Z_N}\sum_{\lambda_i}\langle\lambda_i| \begin{pmatrix} \cos 2\varphi\lambda_+^{j-i} & -\sin 2\varphi\lambda_-^{j-i} \\ -\sin 2\varphi\lambda_+^{j-i} & -\cos 2\varphi\lambda_-^{j-i} \end{pmatrix} \begin{pmatrix} \cos 2\varphi\lambda_+^{N-j+i} & -\sin 2\varphi\lambda_-^{N-j+i} \\ -\sin 2\varphi\lambda_+^{N-j+i} & -\cos 2\varphi\lambda_-^{N-j+i} \end{pmatrix} |\lambda_i\rangle \tag{447}$$

$$= \frac{1}{Z_N}\sum_{\lambda_i}\langle\lambda_i| \begin{pmatrix} \cos^2 2\varphi\lambda_+^N + \sin^2 2\varphi\lambda_+^{N-j+i}\lambda_-^{j-i} & \cdots \\ \cdots & \cos^2 2\varphi\lambda_-^N + \sin^2 2\varphi\lambda_+^{j-i}\lambda_-^{N-j+i} \end{pmatrix} |\lambda_i\rangle \tag{448}$$

$$= \frac{\cos^2 2\varphi(\lambda_+^N + \lambda_-^N) + \sin^2 2\varphi(\lambda_+^{N-j+i}\lambda_-^{j-i} + \lambda_+^{j-i}\lambda_-^{N-j+i})}{\lambda_+^N + \lambda_-^N} \tag{449}$$

Taking the limit $N \to \infty$ gives

$$\langle\sigma_i\sigma_j\rangle = \cos^2 2\varphi + \sin^2 2\varphi\left(\frac{\lambda_-}{\lambda_+}\right)^{j-i} = m^2 + \sin^2 2\varphi\left(\frac{\lambda_-}{\lambda_+}\right)^{j-i}. \tag{450}$$

From the results above, we can easily obtain the correlation function

$$g_{ij} = \langle\sigma_i\sigma_j\rangle - \langle\sigma_i\rangle\langle\sigma_j\rangle = \sin^2 2\varphi\left(\frac{\lambda_-}{\lambda_+}\right)^{j-i}. \tag{451}$$

For the situation $B = 0$, we find that $\cos 2\varphi = 0$ and $\sin 2\varphi = 1$. The values for $\lambda_\pm$ are

$$\lambda_\pm = e^{\beta J} \pm e^{-\beta J} = \begin{cases} 2\cosh\beta J \\ 2\sinh\beta J \end{cases} \tag{452}$$

and the correlation function becomes

$$g_{ij} = \left(\frac{2\sinh\beta J}{2\cosh\beta J}\right)^{j-i} = (\tanh\beta J)^{j-i}. \tag{453}$$

Therefore, for $\beta \to \infty$ (or $T \downarrow 0$), $\langle\sigma_i\sigma_j\rangle$ approaches unity. However, note that for any $T > 0$, $\langle\sigma_i\sigma_j\rangle \to 0$ for $j - i \to \infty$. This is essentially different from below T_c where $\langle\sigma_i\sigma_j\rangle \to m^2$ for $j - i \to \infty$. However, keep in mind that for the one-dimensional Ising model $T_c = 0$.

VI. APPLICATIONS

A. Superfluidity in Helium II

In this section, we will study the unusual properties of liquid ^{4}He (two proton and two neutrons) which is a boson. ^{4}He becomes a liquid at 4.2 K. ^{4}He never becomes a solid at atmospheric pressure, but only a 25 atm. However, around 2.17 K liquid He undergoes another phase transition and part of of the liquid turns into a superfluid. The liquid above 2.17 K is known as He I, whereas below it is known as He II. He I behaves much like an ordinary liquid. Superfluid ^{4}He, on the other hand, has some amazing properties. The viscosity as measured in a flow experiment is essentially zero, and the thermal conductivity is very high. This transition is a directly related to the quantum statistics of a fluid of bosons. A significant number of the helium atoms can be found in the lowest eigenstate, a phenomenon known as Bose-Einstein condensation. The rest of the fluid behaves more or less like a normal fluid. He II can be described with a two-fluid model model. Note that the separation does not occur in real space, as a gas-liquid separation, but it occurs in momentum space. The two components are mixed in real space.

1. Bose-Einstein condensation

Kittel-Kroemer Chapter 7 Let us first discuss the quantum effect that gives rise to the superfluidity. As mentioned above, at a certain temperature we find that a large fraction of the bosons are in the lowest quantum state. This is not entirely surprising since there is no limitation in the number of boson in a particular quantum state. What is surprising is that this occurs not a a temperature of 10^{-14}, the average level spacing but at a temperature of several Kelvin. Let us first have a look at the density of states. The energy of a free boson is given by

$$\varepsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}. \tag{454}$$

Surfaces with a fixed energy are given by the k value. We want to know how many states with have for a given k. The k space taken up by one particular k value is $(\pi/L)^3$. We therefore have to divide a sphere with radius k by this volume

$$N(\varepsilon) = \frac{1}{8}\frac{4}{3}\pi\left(\frac{L}{\pi}\right)^3 k^3 = \frac{V}{6\pi^2}k^3 = \frac{V}{6\pi^2}\left(\frac{2m}{\hbar^2}\varepsilon\right)^{3/2}, \tag{455}$$

where the factor $1/8$ arises from the fact that we only include positive k values and we therefore only need one octant of the sphere. The density of states is given by the change in $N(\varepsilon)$ with respect to energy,

$$\rho(\varepsilon) = \frac{dN(\varepsilon)}{dE} = \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{\varepsilon}. \tag{456}$$

Here we see that having a finite occupation of the lowest state is even more surprising since according to the density of states the lowest state $\varepsilon = 0$ does not contribute at all to the occupation, since

$$N = \int \rho(\varepsilon) f(\varepsilon) d\varepsilon, \tag{457}$$

where $f(\varepsilon)$ is the Bose-Einstein distribution function,

$$f(\varepsilon) = \frac{1}{e^{\beta(\varepsilon-\mu)} - 1}. \tag{458}$$

This is a result of the fact that there are many more quantum states at finite k compared to $k = 0$ of which there is only one. However, replacing the summation by an integral is an approximation that apparently goes horribly wrong here. Let us separate the $k = 0$ term from the integral

$$N = \frac{1}{e^{-\beta\mu} - 1} + \int \rho(\varepsilon) f(\varepsilon) d\varepsilon, \tag{459}$$

where the integral can still start at $\varepsilon = 0$ since $\rho(\varepsilon) = 0$ anyway. The first term gives the occupation of the lowest state. When T approaches zero, the occupancy of the lowest state should approach N, and therefore $\beta\mu \to 0$. This gives

$$N_0 \cong \frac{1}{e^{-\beta\mu} - 1} \cong \frac{1}{1 - \beta\mu - 1} = -\frac{1}{\beta\mu} = -\frac{k_B T}{\mu}. \tag{460}$$

The chemical potential is less than zero. This always has to be the case, since otherwise one would obtain unphysical negative occupation numbers. Note that $\mu = -k_B T/N$ and therefore very small indeed.

Let us now calculate the number of bosons in the states with $k \neq 0$.

$$N_{k\neq 0} = \int \rho(\varepsilon) f(\varepsilon) d\varepsilon = \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty d\varepsilon \frac{\sqrt{\varepsilon}}{e^{\beta(\varepsilon-\mu)}-1}. \tag{461}$$

Since the chemical potential is very close to zero, we can take $e^{-\beta\mu} \cong 1$. Substituing $x = \beta\varepsilon$, we obtain

$$N_{k\neq 0} = \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \frac{1}{\beta^{3/2}} \int_0^\infty dx \frac{\sqrt{x}}{e^x - 1} \tag{462}$$

The integral is evaluated as

$$\int_0^\infty dx \frac{\sqrt{x}}{e^x-1} = \int_0^\infty dx \frac{\sqrt{x}e^{-x}}{1-e^{-x}} = \int_0^\infty dx \sqrt{x} e^{-x} \sum_{n=0}^\infty e^{-nx} = \sum_{n=1}^\infty \int_0^\infty dx \sqrt{x} e^{-nx} = \sum_{n=1}^\infty \frac{1}{s^{3/2}} \int_0^\infty dy \sqrt{y} e^{-y}, \tag{463}$$

using $y = nx$. The sum can be evaluated numerically and gives 2.612. The integral can be rewritten using $y = u^2$ or $u = \sqrt{y}$ giving $dy = 2udu$, and therefore

$$2.612 \times 2 \int_0^\infty du u^2 e^{-u^2} = 2.616 \frac{1}{2}\sqrt{\pi}. \tag{464}$$

Collecting the results gives

$$N_{k\neq 0} = \frac{2.612V}{8}\left(\frac{2mk_BT}{\pi\hbar^2}\right)^{3/2} = 2.612 n_Q V, \tag{465}$$

with $n_Q = (mk_BT/2\pi\hbar^2)^{3/2}$ is the quantum concentration that we found before in the evaluation of an ideal gas. The ratio between the number of electron in $k \neq 0$ and the total number of electrons is therefore

$$\frac{N_{k\neq 0}}{N} = 2.612 \frac{n_Q V}{N} = 2.612 \frac{n_Q}{n}. \tag{466}$$

Einstein condensation temperature. We define the Einstein condensation temperature T_E as the temperature for which the number of atoms in $k \neq 0$ is equal to the total number of atom N. From Eqn. (465), we easily find

$$T_E = \frac{2\pi\hbar^2}{k_B m}\left(\frac{N}{2.612V}\right)^{2/3}. \tag{467}$$

This allows us to rewrite Eqn. (466) as

$$\frac{N_{k\neq 0}}{N} \cong \left(\frac{T}{T_E}\right)^{3/2}. \tag{468}$$

The number of electrons in the ground orbital is then

$$N_{k=0} = N - N_{k\neq 0} = N\left[1 - \left(\frac{T}{T_E}\right)^{3/2}\right]. \tag{469}$$

Note that even for temperatures slightly below the transition temperature, the number of electron in the lowest state is already very large, i.e. of the order of 10^{22} atoms.

The condensation temperature in Kelvin is given by the relation

$$T_E = \frac{115}{V_M^{2/3} M}, \tag{470}$$

where V_M is the molar volume in cm^3/mol and M is the molecular weight. For liquid helium $V_M = 27.6$ cm^3/mol and $M = 4$, giving a $T_E = 3.1$ K. This is remarkably closed to the experimentally observed value of 2.17 K. The interactions between the He atoms are apparently not strong enough to destroy the quantum effects leading to Bose-Einstein condensation. Note that we treated helium here as a gas and not a liquid which is apparently a reasonable approximation.

2. Superfluidity

Kittel and Kroemer, Chapter 7

For many purposes, the superfluid component of helium behaves as a vacuum, as if it were not there. The $N_{k=0}$ electrons in the lowest state have no excitation energy. The superfluid has energy only when the superfluid as a whole is set in motion with respect to the laboratory reference frame.

Let us consider what superfluidity actually means. It implies that the liquid can flow with zero viscosity meaning that the flow does not create excitations from the $k = 0$ states to the $k \neq 0$ states. These transitions could occur by imperfections at the wall or other impurities. An alternative way to look at the same thing is that an object or a particle (say, a neutron) can more through the superfluid without feeling any damping whatsoever. Let us consider the conditions to create elementary excitations in He II with a specific energy $\varepsilon_{\mathbf{k}}$ and momentum $\hbar\mathbf{k}$. To create such an excitation we require conservation of energy

$$\frac{1}{2}m_0 v^2 = \frac{1}{2}m_0 v'^2 + \varepsilon_{\mathbf{k}}, \tag{471}$$

and momentum

$$m_0\mathbf{v} = m_0\mathbf{v}' + \hbar\mathbf{k}, \tag{472}$$

where m_0 is the mass of the object moving through the superfluid. Let us rewrite the latter equation

$$m_0\mathbf{v} - \hbar\mathbf{k} = m_0\mathbf{v}'. \tag{473}$$

Squaring gives

$$m_0^2\mathbf{v}^2 - \hbar m_0\mathbf{v}\cdot\mathbf{k} + \hbar^2\mathbf{k}^2 = m_0^2\mathbf{v}'^2. \tag{474}$$

Multiplying by $\frac{1}{2m_0}$ gives

$$\frac{1}{2}m_0\mathbf{v}^2 - \frac{1}{2}\hbar\mathbf{v}\cdot\mathbf{k} + \frac{1}{2m_0}\hbar^2\mathbf{k}^2 = \frac{1}{2}m_0\mathbf{v}'^2. \tag{475}$$

Subtracting this from the equation for conservation of energy gives

$$\frac{1}{2}\hbar\mathbf{v}\cdot\mathbf{k} - \frac{1}{2m_0}\hbar^2\mathbf{k}^2 = \varepsilon_{\mathbf{k}}. \tag{476}$$

We then find for the velocity

$$v = \frac{\varepsilon_{\mathbf{k}} + \frac{1}{2m_0}\hbar^2\mathbf{k}^2}{\hbar k\cos\varphi}, \tag{477}$$

where φ is the angle between $\mathbf{v}$ and $\mathbf{k}$. We are looking for the maximum allowable velocity that creates no excitations. This is given by the minimum of the right-hand side. The minimum is given by $\varphi = 0$. Let us take the mass of the object $m_0 \to \infty$. This gives

$$v = \text{Minimum } \frac{\varepsilon_{\mathbf{k}}}{\hbar k}. \tag{478}$$

The only excitations that we have are He atoms that have an energy $\varepsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}$, so we have

$$v = \text{Minimum } \frac{\hbar k}{2m}. \tag{479}$$

Apparently, we have a problem here since this minimum is zero. This seems to imply that there is no superfluidity, unless we do not understand the elementary excitations in superfluid He II. What we need is something with a linear dispersion, say $\varepsilon_k = \hbar\omega = \hbar v_s k$. This would give for the velocity

$$v = \text{Minimum } \frac{\hbar v_s k}{\hbar k} = v_s. \tag{480}$$

Pathria Chapter 10
In the previous section, we saw that independent He atoms do not give superfluidity. Apparently, we missed something. What we need to include are the interaction between the ^{4}He atoms. The interaction $u(\mathbf{r}_2-\mathbf{r}_1)$ will only depend on the relative positions of two helium atoms. The Hamiltonian is given by

$$H=\sum_{\mathbf{k}}\varepsilon_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}+\sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}U_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4}a_{\mathbf{k}_4}^{\dagger}a_{\mathbf{k}_3}^{\dagger}a_{\mathbf{k}_2}a_{\mathbf{k}_1}. \tag{481}$$

The interaction matrix element is given by

$$U_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4}=\int d\mathbf{r}_1\int d\mathbf{r}_2\varphi_{\mathbf{k}_4}^{*}(\mathbf{r}_1)\varphi_{\mathbf{k}_3}^{*}(\mathbf{r}_1)u(\mathbf{r}_2-\mathbf{r}_1)\varphi_{\mathbf{k}_2}(\mathbf{r}_2)\varphi_{\mathbf{k}_1}(\mathbf{r}_1). \tag{482}$$

If we take the basis set to be plane waves, we obtain

$$U_{\mathbf{k}_1,\mathbf{k}_2a,\mathbf{k}_3,\mathbf{k}_4}=\frac{1}{V^2}\int d\mathbf{r}_1\int d\mathbf{r}_2e^{-i(\mathbf{k}_3-\mathbf{k}_2)\cdot\mathbf{r}_2}u(\mathbf{r}_2-\mathbf{r}_1)e^{-i(\mathbf{k}_4-\mathbf{k}_1)\cdot\mathbf{r}_1}. \tag{483}$$

Putting $\mathbf{r}=\mathbf{r}_2-\mathbf{r}_1$, we can rewrite this as

$$U_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4}=\frac{1}{V^2}\int d\mathbf{r}\int d\mathbf{r}_2e^{-i(\mathbf{k}_3-\mathbf{k}_2+\mathbf{k}_4-\mathbf{k}_1)\cdot\mathbf{r}_2}u(\mathbf{r})e^{i(\mathbf{k}_4-\mathbf{k}_1)\cdot\mathbf{r}}. \tag{484}$$

This integral contains a delta function

$$\frac{1}{V}\int d\mathbf{r}_2e^{-i(\mathbf{k}_3-\mathbf{k}_2+\mathbf{k}_4-\mathbf{k}_1)\cdot\mathbf{r}_2}=\delta(\mathbf{k}_3-\mathbf{k}_2+\mathbf{k}_4-\mathbf{k}_1)\quad\Rightarrow\quad\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}_4+\mathbf{k}_3. \tag{485}$$

We can always satisfy this by taking $\mathbf{k}_1=\mathbf{k}$, $\mathbf{k}_2=\mathbf{k}$, $\mathbf{k}_3=\mathbf{k}-\mathbf{q}$, and $\mathbf{k}_4=\mathbf{k}+\mathbf{q}$. We see that there is only one free wavevector for the interaction $U_{\mathbf{k}_1,\mathbf{k}_2a,\mathbf{k}_3,\mathbf{k}_4}\rightarrow U_{\mathbf{q}}$. Defining

$$U_{\mathbf{q}}=\frac{1}{V}\int d\mathbf{r}u(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}, \tag{486}$$

we can write the Hamiltonian as

$$H=\sum_{\mathbf{k}}\varepsilon_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}+\sum_{\mathbf{k}\mathbf{k}'\mathbf{q}}U_{\mathbf{q}}a_{\mathbf{k}+\mathbf{q}}^{\dagger}a_{\mathbf{k}'-\mathbf{q}}^{\dagger}a_{\mathbf{k}'}a_{\mathbf{k}}. \tag{487}$$

We want to use this Hamiltonian to describe the collisions between the helium atoms at low temperature. Since superfluidity occurs at a few Kelvin and less, we are dealing with small momenta and we can take

$$U_{\mathbf{q}=0}=\frac{1}{V}\int d\mathbf{r}u(\mathbf{r})=\frac{u_0}{V}. \tag{488}$$

The Hamiltonian in this limit is

$$H\cong\sum_{\mathbf{k}}\varepsilon_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}+\frac{u_0}{V}\sum_{\mathbf{k}\mathbf{k}'}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}'}^{\dagger}a_{\mathbf{k}'}a_{\mathbf{k}}. \tag{489}$$

Let us split this up into scattering into the same momentum and into different momenta

$$H\cong\sum_{\mathbf{k}}\varepsilon_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}+\frac{u_0}{V}\sum_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}a_{\mathbf{k}}+\frac{u_0}{V}\sum_{\mathbf{k}\neq\mathbf{k}'}\left[a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}'}^{\dagger}a_{\mathbf{k}'}a_{\mathbf{k}}+a_{\mathbf{k}'}^{\dagger}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}'}a_{\mathbf{k}}\right]. \tag{490}$$

Let us evaluate the different components

$$\frac{u_0}{V}\sum_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}a_{\mathbf{k}}=\frac{u_0}{V}\sum_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}(a_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}-1)a_{\mathbf{k}}=\frac{u_0}{V}\sum_{\mathbf{k}}(n_{\mathbf{k}}^2-n_{\mathbf{k}})=\frac{u_0}{V}\sum_{\mathbf{k}}(n_{\mathbf{k}}^2-N). \tag{491}$$

The other terms are

$$\frac{u_0}{V}\sum_{\mathbf{k}\neq\mathbf{k}'} a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}'} a_{\mathbf{k}'} a_{\mathbf{k}} = \frac{u_0}{V}\sum_{\mathbf{k}\neq\mathbf{k}'} n_{\mathbf{k}} n_{\mathbf{k}'} = \frac{u_0}{V}\sum_{\mathbf{k}} n_{\mathbf{k}} \sum_{\mathbf{k}'\neq\mathbf{k}} n_{\mathbf{k}'} = \frac{u_0}{V}\sum_{\mathbf{k}} n_{\mathbf{k}}(N - n_{\mathbf{k}}) = N^2 - \sum_{\mathbf{k}} n^2_{\mathbf{k}}. \tag{492}$$

The exchange term gives the same result. The energy of the system can then be written as

$$E = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}} + \frac{u_0}{V}\left[2N^2 - N - \sum_{\mathbf{k}} n^2_{\mathbf{k}}\right] \tag{493}$$

$$\cong \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}} + \frac{u_0}{V}\left[2N^2 - n_0^2\right]. \tag{494}$$

Let us examine the ground state which approximately corresponds to

$$n_{\mathbf{k}} \cong \begin{cases} N & \text{for } \mathbf{k} = 0 \\ 0 & \text{for } \mathbf{k} \neq 0 \end{cases}, \tag{495}$$

giving for the energy

$$E = \frac{u_0 N^2}{V}. \tag{496}$$

The ground-state pressure is

$$P = -\left(\frac{\partial E}{\partial V}\right)_N = \frac{u_0 N^2}{V^2} = u_0 n^2, \tag{497}$$

where n is the particle density. From the pressure, we can obtain the sound velocity through

$$v_s = \frac{1}{m}\frac{dP}{dn} = \frac{2u_0 n}{m}. \tag{498}$$

Apparently, we do have sound waves in the ground state. This gives us hope that we might indeed have superfluidity in He II. Let us now evaluate the low-lying excited states of He with interactions.

4. Elementary Excitations in He II

We will now evaluate the low-lying states in an imperfect boson gas, namely He in the presence of interactions between the atoms. To keep things simple, we assume that most of the atoms are in the state $k = 0$, and that the occupation of the states with $k \neq 0$ is small. Let us go back to Eqn. (494). We can rewrite the interaction energy using

$$2N^2 - n_0^2 = N^2 + (N^2 - n_0^2) = N^2 + (N + n_0)(N - n_0) \cong N^2 + 2N(N - n_0) = N^2 + 2N\sum_{\mathbf{k}\neq 0} a^\dagger_{\mathbf{k}} a_{\mathbf{k}}. \tag{499}$$

using the fact that $n_0 \cong N$ We need to add an additional term namely the scattering between the states with $k = 0$ to the states with finite momentum

$$\sum_{\mathbf{k}\neq 0} U(\mathbf{k})\left[a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a_0 a_0 + a^\dagger_0 a^\dagger_0 a_{\mathbf{k}} a_{-\mathbf{k}}\right]. \tag{500}$$

Note that pairs of He atoms in the excited state are created with opposite momentum since the total momentum has to remain zero. We would now like to replace the a_0 and $a_0^\dagger$ operators by numbers. We know that $a_0^\dagger a_0 = n_0 \to N$. We also have from the commutation relation $a_0 a_0^\dagger = a_0^\dagger a_0 + 1 \cong a_0^\dagger a_0 \to N$. We therefore assume that both a_0 and $a_0^\dagger$ give a numberclose to $\sqrt{N}$. Furthermore, since we are dealing only with low-lying excitations, we can replace $U(\mathbf{k}) = \frac{u_0}{V}$. This gives for the total Hamiltonian

$$H = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{u_0}{V}\left[N^2 + N\sum_{\mathbf{k}\neq 0}(2a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} + a_{\mathbf{k}} a_{-\mathbf{k}})\right], \tag{501}$$

or

$$H = \sum_{\mathbf{k}} (\varepsilon_{\mathbf{k}} + 2u_0 n) a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + u_0 n \sum_{\mathbf{k}\neq 0} (a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} + a_{\mathbf{k}} a_{-\mathbf{k}}) + u_0 n^2 V \tag{502}$$

where $n = N/V$.

Intermezzo: displaced harmonic oscillator. We would like to diagonalize the Hamiltonian above. The problem is somewhat comparable to a displaced harmonic oscillator

$$H = \hbar\omega a^\dagger a + K(a^\dagger + a). \tag{503}$$

This can be diagonalized by introducing displaced oscillators $b^\dagger = a^\dagger + \Delta$. The diagonalized Hamiltonian can be written as

$$H = \hbar\omega b^\dagger b + E_0 = \hbar\omega a^\dagger a + \hbar\omega\Delta(b^\dagger + b) + \hbar\omega\Delta^2 + E_0, \tag{504}$$

with $\Delta = \frac{K}{\hbar\omega}$ and $E_0 = -\frac{K^2}{\hbar\omega}$ this Hamiltonian becomes equivalent to the one given above.

In the spirit of the displaced harmonic oscillator, we introduce new operators

$$b^\dagger_{\mathbf{k}} = \cosh x a^\dagger_{\mathbf{k}} + \sinh x a_{-\mathbf{k}} \quad \text{and} \quad b_{\mathbf{k}} = \cosh x a_{\mathbf{k}} + \sinh x a^\dagger_{-\mathbf{k}}. \tag{505}$$

The advantage of using the hyperbolic functions is apparent we we look at the commutation relationships

$$\begin{aligned} [b_{\mathbf{k}}, b^\dagger_{\mathbf{k}}] &= (\cosh x a_{\mathbf{k}} + \sinh x a^\dagger_{-\mathbf{k}})(\cosh x a^\dagger_{\mathbf{k}} + \sinh x a_{-\mathbf{k}}) - (\cosh x a^\dagger_{\mathbf{k}} + \sinh x a_{-\mathbf{k}})(\cosh x a_{\mathbf{k}} + \sinh x a^\dagger_{-\mathbf{k}}) & (506)\\ &= \cosh^2 x [a_{\mathbf{k}}, a^\dagger_{\mathbf{k}}] + \sinh^2 x [a^\dagger_{-\mathbf{k}}, a_{-\mathbf{k}}] = (\cosh^2 x - \sinh^2 x) = 1 & (507) \end{aligned}$$

where we used the fact that $[a_{\mathbf{k}}, a^\dagger_{\mathbf{k}}] = 1$, $[a^\dagger_{\mathbf{k}}, a^\dagger_{\mathbf{k}}] = 0$, and $[a_{\mathbf{k}}, a_{\mathbf{k}}] = 0$; we also used the relationship $\cosh^2 x - \sinh^2 x = 1$. The operators $b_{\mathbf{k}}$ and $b^\dagger_{\mathbf{k}}$ are therefore proper boson operators. We would now like to write our Hamiltonian in diagonal form

$$H = \sum_{\mathbf{k}} E_{\mathbf{k}} b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + E_0. \tag{508}$$

This can be written as

$$H = \sum_{\mathbf{k}} \left[E_{\mathbf{k}} \cosh^2 x a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + E_{\mathbf{k}} \sinh^2 x a_{-\mathbf{k}} a^\dagger_{-\mathbf{k}} + E_{\mathbf{k}} \sinh x \cosh x (a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} + a_{-\mathbf{k}} a_{\mathbf{k}}) \right] + E_0, \tag{509}$$

which gives

$$H = \sum_{\mathbf{k}} \left[E_{\mathbf{k}} \cosh 2x a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2} E_{\mathbf{k}} \sinh 2x (a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} + a_{\mathbf{k}} a_{-\mathbf{k}}) \right] + \sum_{\mathbf{k}} E_{\mathbf{k}} \sinh^2 x + E_0, \tag{510}$$

where we have used the commutation relationship and the fact that summations over $\mathbf{k}$ can be replaced by summations over $-\mathbf{k}$. We also used the relationships

$$\cosh 2x = \cosh^2 x + \sinh^2 x \quad \text{and} \quad \sinh 2x = 2 \sinh x \cosh x \tag{511}$$

This Hamiltonian has to be equivalent to that in Eqn. (502). For this we require that

$$\begin{aligned} E_{\mathbf{k}} \cosh 2x &= \varepsilon_{\mathbf{k}} + 2u_0 n \quad \Rightarrow \quad \cosh 2x = \frac{\varepsilon_{\mathbf{k}} + 2u_0 n}{E_{\mathbf{k}}} & (512)\\ \frac{1}{2} E_{\mathbf{k}} \sinh 2x &= u_0 n \quad \Rightarrow \quad \sinh 2x = \frac{2u_0 n}{E_{\mathbf{k}}} & (513) \end{aligned}$$

Dividing the two expressions gives

$$\tanh 2x = \frac{2u_0 n}{\varepsilon_{\mathbf{k}} + 2u_0 n}. \tag{514}$$

We also have

$$\cosh^2 2x - \sinh^2 2x \;=\; 1 \quad \Rightarrow \quad \left(\frac{\varepsilon_{\mathbf{k}} + 2u_0 n}{E_{\mathbf{k}}}\right)^2 - (\frac{2u_0 n}{E_{\mathbf{k}}})^2 = 1 \tag{515}$$

giving

$$E_{\mathbf{k}} = \sqrt{(\varepsilon_{\mathbf{k}} + 2u_0 n)^2 - (2u_0 n)^2} = \sqrt{\varepsilon_{\mathbf{k}}^2 + 4u_0 n \varepsilon_{\mathbf{k}}}. \tag{516}$$

Often these expression are found somewhat differently following the original work by Bogoliubov. Using the formula for the double angle of the tanh, we can rewrite Eqn. (514)

$$\frac{2\tanh x}{1+\tanh^2 x} = \frac{2\alpha_{\mathbf{k}}}{1+\alpha_{\mathbf{k}}^2} = \frac{2u_0 n}{\varepsilon_{\mathbf{k}} + 2u_0 n}. \tag{517}$$

where we define $\alpha_{\mathbf{k}} \equiv \tanh x$. This leads to a quadratic equation for $\alpha_{\mathbf{k}}$

$$u_0 n \alpha_{\mathbf{k}}^2 - (\varepsilon_{\mathbf{k}} + 2u_0 n)\alpha_{\mathbf{k}} + u_0 n = 0, \tag{518}$$

or

$$\alpha_{\mathbf{k}} = \frac{1}{2u_0 n}\left[\varepsilon_{\mathbf{k}} + 2u_0 n - \sqrt{(\varepsilon_{\mathbf{k}} + 2u_0 n) - 4(u_0 n)^2}\right] = \frac{1}{2u_0 n}\left[\varepsilon_{\mathbf{k}} + 2u_0 n - E_{\mathbf{k}}\right]. \tag{519}$$

Since we have

$$\alpha_{\mathbf{k}} = \tanh x = \frac{\sinh x}{\cosh x} \quad \Rightarrow \quad \sinh x = \frac{\alpha_{\mathbf{k}}}{\Delta} \quad \text{and} \quad \cosh x = \frac{1}{\Delta}. \tag{520}$$

The value for Δ can be determined from

$$\cosh^2 x - \sinh^2 x = \frac{1}{\Delta^2} - \frac{\alpha_{\mathbf{k}}^2}{\Delta^2} = 1 \quad \Rightarrow \quad \Delta = \sqrt{1-\alpha_{\mathbf{k}}^2}. \tag{521}$$

Using that in the expressions for the $\sinh x$ and $\cosh x$ we arrive at the traditional Bogoliubov operators

$$b_{\mathbf{k}}^\dagger = \frac{a_{\mathbf{k}}^\dagger + \alpha_{\mathbf{k}} a_{-\mathbf{k}}}{\sqrt{1-\alpha_{\mathbf{k}}^2}} \quad \text{and} \quad b_{\mathbf{k}} = \frac{a_{\mathbf{k}} + \alpha_{\mathbf{k}} a_{-\mathbf{k}}^\dagger}{\sqrt{1-\alpha_{\mathbf{k}}^2}}. \tag{522}$$

Usually, one includes in the calculation of E_0 also higher-order terms in $U(\mathbf{k})$ we shall not consider here. We end up with the diagonal Hamiltonian

$$H = \sum_{\mathbf{k}} E_{\mathbf{k}} b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + E_0 = \sum_{\mathbf{k}} \sqrt{\left(\frac{\hbar^2}{2m}\right)^2 k^4 + \frac{2u_0 n \hbar^2}{m} k^2} b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + E_0 \tag{523}$$

For small k, $E_{\mathbf{k}} \sim k$ and we have a linear dispersion in k. This means that we can expect superfluidity.

Potential.– In this section, some very handwaving arguments are given about the potential in the previous section. The interactions between He atoms are given by a Lennard-Jones type potential

$$u(r) = U_0 \left\{ \left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^6 \right\}. \tag{524}$$

This potential is repulsive at short distances and has a small attractive interaction. Unfortunately, this potential is hard to evaluate, so let us simplify this enormously. Let us take a potential

$$u(r) = \begin{cases} U & r < a_0 \\ 0 & r > a_0 \end{cases} \tag{525}$$

This interaction takes into account the strong repulsive part around the He atom and neglects all the other r dependence. It is slightly unphysical since the potential does not go to infinity, but it greatly simplifies the calculations.

We now want to find the Fourier transform of the potential, see Eqn. (matrixU). When $u(\mathbf{r})$ only depends on r, we can rewrite the matrix element. Taking $\mathbf{q}$ along the z axis gives

$$U_{\mathbf{q}} = \frac{1}{V}\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \int_0^{\infty} drr^2 \sin\theta u(r) e^{iq\cos\theta} = 2\pi \int_{-1}^{1} d(\cos\theta) \int_0^{\infty} drr^2 u(r) e^{iqr\cos\theta} \tag{526}$$

$$= 2\pi \int_0^{\infty} drr^2 u(r) \left[\frac{1}{iqr} e^{iqr\cos\theta}\right]_{-1}^{1} = 4\pi \int_0^{\infty} drr^2 u(r) \frac{1}{qr} \frac{e^{iqr} - e^{-iqr}}{2i}. \tag{527}$$

This gives for a central potential

$$u_{\mathbf{q}} = 4\pi \int_0^{\infty} drr^2 u(r) \frac{\sin qr}{qr}. \tag{528}$$

This integral simplifies when using our approximate potential

$$u_{\mathbf{q}} = 4\pi U \int_0^{a_0} r \frac{\sin qr}{q}. \tag{529}$$

We can solve this using partial integration. Taking $v = r$ and $u' = \sin qr$ gives $v' = 1$ and $u = -\frac{1}{q}\cos qr$. Partial integration gives

$$u_{\mathbf{q}} = \frac{4\pi U}{q^2}\left\{-[r\cos qr]_0^{a_0} + \int_0^{a_0} dr \cos qr\right\} = \frac{4\pi U}{q^2}\left\{-a_0\cos qa_0 + \frac{1}{q}[\sin qr]_0^{a_0}\right\}, \tag{530}$$

yielding

$$u_{\mathbf{q}} = \frac{4\pi U a_0}{q^2}\left\{\frac{\sin qa_0}{qa_0} - \cos qa_0\right\}. \tag{531}$$

In the limit $q \to 0$, this gives

$$u_{\mathbf{q}} = \frac{4\pi U a_0}{q^2}\left\{1 - (1 - \frac{1}{2}(qa_0)^2\right\} = \frac{4\pi U a_0^3}{2} - \mathcal{O}(q^2). \tag{532}$$

This is a constant which is exactly what we need to obtain a linear dispersion in the small q limit.

Homework 4/28/2007
For questions veenendaal@niu.edu

Problem 1
Let us consider N three-dimensional harmonic oscillators

$$H = \sum_{i=1}^{N}(n_i + \frac{3}{2})\hbar\omega, \tag{533}$$

where $n_i = n_{i,x} + n_{i,y} + n_{i,z} = 0, 1, 2, 3, \cdots$. The partition function is given by

$$Z_N = \prod_{i=1}^{N} Z_i = \prod_{i=1}^{N} \sum_{n_i=0}^{\infty} g(n_i) e^{-\beta E_i(n_i)}, \tag{534}$$

where $E_i(n_i) = (n_i + \frac{3}{2})\hbar\omega$ and $g(n_i)$ is the degeneracy of the energy level labelled by n_i.

(a) Show that

$$g(n_i) = \frac{(n_i + 2)!}{2! n_i!} \tag{535}$$

and express $g(n_i)$ as a polynomial. Hint: this is essentially a problem of dividing n_i items in three boxes.
(b) We want to sum over n_i. For this we need

$$\sum_{n=0}^{\infty} n q^n = \frac{q}{(1-q)^2} \tag{536}$$

$$\sum_{n=0}^{\infty} n^2 q^n = \frac{q(1+q)}{(1-q)^3}, \tag{537}$$

where $q = \exp(-\beta\hbar\omega)$. Show that the first of these expressions can be derived by taking the derivative with respect to q of $\sum_{n=0}^{\infty} q^n$ (we just assume the second one to be correct).
(c) Calculate the partition function and show that it can also be written as

$$Z_N = \left[\frac{e^{-\frac{1}{2}\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}\right]^{3N} \tag{538}$$

(d) Give an alternative (and much shorter) derivation of this result and explain why this can be derived in this way.
(e) Show that the energy can be written as

$$E = \frac{3}{2} N\hbar\omega \coth \frac{\hbar\omega}{2k_B T} \tag{539}$$

Problem 2
The estimate of the configurational partition function (i.e. the part of the partition function related to the integration over the volume) for a gas of hard sphere molecules (of diameter σ) may be obtained in the following way. If there are two molecules, one can be anywhere in the vessel and the other is restricted to a volume $V - \frac{4}{3}\pi\sigma^3$. If there are s molecules in the gas, we may suppose that each molecule restricts the volume available to the others and write

$$Q_N \cong V(V - v_0)(V - 2v_0)(V - 3v_0)\cdots(V - (s-1)v_0), \tag{540}$$

where $v_0 = \frac{4}{3}\pi\sigma^3$.
(a) Explain why $v_0 = \frac{4}{3}\pi\sigma^3$ (note that σ is the diameter of the hard sphere molecule and not the radius).
(b) Show that $Q_N \cong (V - \frac{1}{2}N v_0)^N$ if $V > N\frac{4}{3}\pi\sigma^3$.
Hint: Express the product in the small parameter v_0/V and then turn the product into a sum by taking the logarithm.