Statistical Physics

Yuri Galperin and Jens Feder

FYS 203

The latest version at www.uio.no/ yurig

Contents

Chapter 1

Introduction

1.1 Motivation

Statistical physics is an unfinished and highly active part of physics. We do not have the theoretical framework in which to even attempt to describe highly irreversible processes such as fracture. Many types of nonlinear systems that lead to complicated pattern formation processes, the properties of granular media, earthquakes, friction and many other systems are beyond our present understanding and theoretical tools.

In the study of these and other macroscopic systems we build on the established conceptual framework of *thermodynamics* and *statistical physics*. Thermodynamics, put into its formal and phenomenological form by Clausius, is a theory of impressive range of validity. Thermodynamics describes all systems form classical gases and liquids, through quantum systems such as superconductors and nuclear matter, to black holes and elementary particles in the early universe in exactly the same form as they were originally formulated.[1]

Statistical physics, on the other hand gives a rational understanding of Thermodynamics in terms of microscopic particles and their interactions. Statistical physics allows not only the calculation of the temperature dependence of thermodynamics quantities, such as the specific heats of solids for instance, but also of transport properties, the conduction of heat and electricity for example. Moreover, statistical physics in its modern form has given us a complete understanding of second-order phase transitions, and with Wilson's *Renormalization Group theory* we may calculate the scaling exponents observed in experiments on phase transitions.

However, the field of statistical physics is in a phase of rapid change. New ideas and concepts permit a fresh approach to old problems. With new concepts we look for features ignored in previous experiments and find exciting results. Key words are *deterministic chaos, fractals, self-organized criticality (SOC), turbulence* and *intermitency*. These words represent large fields of study which have changed how we view Nature.

Disordered systems, percolation theory and fractals find applications not only in physics and engineering but also in economics and other social sciences.

What are the processes that describe the evolution of *complex systems*? Again we have an

---

[1] From the introduction of Ravndal's lecture notes on Statistical Physics

active field with *stochastic processes* as a key word.

So what has been going on? Why have the mysteries and excitement of statistical physics not been introduced to you before?

The point is that statistical physics is unfinished, subtle, intellectually and mathematically demanding. The activity in statistical physics has philosophical implications that have not been discussed by professional philosophers—and have had little influence on the thinking of the general public.

DETERMINISTIC behavior rules the day. Privately, on the public arena, in industry, management and in teaching we almost invariably assume that any given action leads predictably to the desired result—and we are surprised and discouraged when it does not.[2] The next level is that some feed-back is considered—but again things get too simplistic. Engineers who study road traffic reach the conclusion that more roads will make congestions worse, not better—a statement that cannot be generally true. Similarly, the "Club of Rome" report predicted that we would run out of resources very quickly. The human population would grow to an extent that would outstrip food, metals, oil and other resources. However, this did not happen. Now, with a much large population we have more food, metals, and oil per capita than we had when the report was written. Of course, statistical physics cannot be applied to all areas. The point is that in physics, statistical physics provides us with many examples of complex systems and complex dynamics that we may understand in detail.

Statistical physics provides an intellectual framework and a systematic approach to the study of real world systems with complicated interactions and feedback mechanisms. I expect that the new concepts in statistical physics will eventually have a significant impact not only on other sciences, but also on the public sector to an extent that we may speak of as a paradigm shift.

1.2 Deterministic Chaos

Deterministic chaos appears to be an *oxymoron*, i. e. a contradiction of terms. It is not. The concept of deterministic chaos has changed the thinking of anyone who gets some insight into what it means. Deterministic chaos may arise in any non-linear system with few (active) degrees of freedom, e. g. electric oscillator circuits, driven pendulums, thermal convection, lasers, and many other systems. The central theme is that the time evolution of the system is *sensitive* so initial conditions. Consider a system that evolves in time as described by the deterministic equations of motion, a pendulum driven with a small force oscillating at a frequency ω, for example. Starting from given initial condition the pendulum will evolve in time as determined by the equation of motion. If we start again with new initial condition, arbitrarily close to the previous case, the pendulum will again evolve deterministically, but its state as a function of time will have a distance (in some measure) that diverges exponentially from the first realization. Since the initial condition cannot be specified with arbitrary precision, the orbit in fact becomes unpredictable—even in principle.

It is this incredible sensitivity to initial conditions for *deterministic* systems that has deep

---

[2] Discuss the telephone system at the University of Oslo

philosophical implications. The conclusion, at least how I see it, is that determinism and randomness are just two aspects of same system.

The planetary system is known to be chaotic! That is, we cannot, even in principle, predict lunar eclipse, the relative positions of planets, the rotation of the earth as a function of time. What is going on? The point is that non-linear systems that exhibit deterministic chaos have a *time horizon* beyond which we cannot predict the time evolution because of the exponential sensitivity on initial conditions. For the weather we have a time horizon of two weeks in typical situations. There exist initial conditions that have time horizons further in the future. for the planetary system the time horizon is millions of years, but chaotic behavior has nevertheless been observed for the motion of Io, one of Jupiter's moons. Chaos is also required for an understanding of the banding of the rings of Jupiter.[3] Very accurate numerical solutions of Newton's equations for the solar system also exhibit deterministic chaos.

1.3 Atomistic Philosophy

Now, for systems with many active degrees of freedom, we observe *cooperative phenomena*, i. e. behavior that does not depend on the details of the system. We observe *universality*, or universal behavior, that cannot be explained by an atomistic understanding alone.

Let me explain: The *atomistic philosophy* is strong and alive. We believe that with an atomistic understanding of the particles that constitute matter—and their interactions—it is merely an exercise in applied mathematics to predict the macroscopic behavior of systems of practical importance. This certainly is an arrogant physicists attitude. It must be said, however, that the experiments and theories are entirely consistent with this view—in spite of deterministic chaos, self-organized criticality and their ilk.

It is clear that the quantum mechanics required to describe the interactions of atoms in chemical systems is fully understood. We have the exact solution for the Hydrogen atom, the approximate solution for the Helium atom, and numerical solutions for more complex atoms and their interactions in molecules. There is no evidence that there are open questions in the foundation of quantum chemistry. Thus we know what is required for a rational understanding of chemistry—the work has been done, we only need to apply what is known practical tasks in engineering and industry so to say.

Since we understand chemistry, no new fundamental laws of Nature are required to understand biochemistry, and finally by implication the brain—how we think and understand!

What folly! Clearly something is wrong with this line of reasoning. Molecular biology is at present the field of science that has the highest rate of discovery, while physicist have been waiting for almost thirty years for the observation of the Higgs Boson predicted by the standard model. Elementary particle theory clearly lies at the foundation of physics—but the rate of discovery has slowed to a glacial pace.

What is going on? I believe the point is that the evolution of science in the 20th century has lead to a focus on the atomistic part of the atomistic philosophy. Quantum mechanics was a

---

[3] The ring were first observed by Galileo in 1616 with a telescope he had built. By 1655 Huygens had resolved the banding, not fully explained yet.

fantastic break-through, which completely dominated physics in the first half of previous century. Discoveries, were made at an astounding rate. Suddenly we understood spectra, the black-body radiation, X-rays, radioactivity, all kinds of new particles, electrons, protons, neutrons, anti-particles and much more. Much was finished before the second World war, whereas nuclear reactors, and weapons, were largely developed in 50's and 60's. The search for elementary particles, the constituents of nuclear matter, and their interactions gained momentum in the 50's and truly fantastic research organizations of unprecedented size, such as CERN, dedicated to basic research were set up.

This enthusiasm was appropriate and lead to a long string of discoveries (and Nobel prizes) that have changed our society in fundamental ways.

Quantum mechanics is, of course, also at the basis of the electronics industry, computers, communication, lasers and other engineering products that have changed our lives.

However, this huge effort on the atomistic side, left small resources—both human and otherwise—to other parts of physics. The intellectual resources were largely concentrated on particle physics research, the other aspects of the atomistic philosophy, namely the understanding of the natural world in terms of the (newly gained) atomistic physics, was left to an uncoordinated little publicized and conceptually demanding effort.

1.3.1 Cooperative Phenomena

Cooperative phenomena are ubiquitous and well known from daily life. Take hydrodynamics for example. Air, water, wine, molten steel, liquid argon — all share the same hydrodynamics. They flow, fill vessels, they from drops and bubbles, have surface waves, vortices and exhibit many other hydrodynamic phenomena—in spite of the fact that molten steel, water and argon are definitely described by quite different Hamilton operators that encode their quantum mechanics. How can it be that important macroscopic properties do not depend on the details of the Hamiltonian?

Phase transitions are also *universal*: Magnetism and the liquid–vapor transition belong to the same universality class; in an abstract sense they are the same phase transition! Ferro-electricity, spinodal decomposition, superconductivity, are all phenomena that do not depend on the details of their atomistic components and their interactions—they are cooperative phenomena.

The understanding of cooperative phenomena is far from complete. We have no general theory, except for second order phase transitions, and the phenomenological equations for hydrodynamics, to classify and simplify the general scaling and other fractal phenomena observed in highly non-equilibrium systems. Here we have a very active branch of statistical physics.

1.4 Outlook

So what is the underlying conceptual framework on which physics attempts to fill the gap in understanding the macroscopic world? How do we connect the microscopic with the macroscopic?—of course, by the scientific approach, by observation, experiment, theory, and now computational modeling.

First of all: For macroscopic systems, we have a fantastic *phenomenological theory*: THERMODYNAMICS valid at or near thermal equilibrium, and for systems that start in an equilibrium state and wind up in an equilibrium state (explosions for example). Thermodynamics was developed largely in the 19th century, with significant developments in our century related to quantum mechanics and phase transitions. Thermodynamics in the present form was really formulated as an axiomatic system with the three laws of thermodynamics. The central concept is energy[4], and of course *entropy*, the only concept required in all the three laws of thermodynamics.

Secondly: Statistical physics started with Daniel Bernoulli (1700-1792), was given a new start by Rudolf Clausius (1822–1888), James Clerk Maxwell (1831–1879) contributed the kinetic theory of gases and his famous velocity distribution. Ludwig Boltzmann (1844–1906) made the fundamental connection to kinetics and introduced the famous expression $S = k \ln W$ the statistical expression for the entropy, and explained why entropy must increase by his H-theorem. Josiah Willard Gibbs (1839–1903), made fundamental contributions to the modern formulation of statistical mechanics. In this century Lars Onsager (1903–1976) made several outstanding contributions. By his exact solution of the Ising model, in two-spatial dimensions, he proved that the framework of statistical physics could indeed tackle the problem of phase transitions. He received the 1968 Nobel prize in chemistry for his work on irreversible thermodynamics. Claude E. Shannon initiated *information theory* by his 1948 paper analyzing measures of information transmitted through a communication channel. His measure of information is directly related to Boltzmann's statistical measure of entropy. The last break-trough contribution was by Kenneth Geddes Wilson (1936–), who was awarded the 1982 Nobel prize in physics for Renormalization group theory that allows the calculation of scaling exponents at second order phase transitions.

1.5 What will we do?

This is what now think we should do in this course. We cannot get into all the interesting things at the frontier of research. On the contrary the purpose of this course is to acquaint you with the central issues of statistical mechanics. I will be guided by Finn Ravndals notes, and my own tastes, and hope I do not overload the course. It is better to understand the a number of central issues thoroughly, they provide anchor points for further investigation if your studies, or later work require them.

---

[4] Energy, an old word first used by Thomas Young in 1807 to signify mv^2, not $[\frac{1}{2}mv^2]$ from Greek *en* in +*ergon* work (or rather: at work)

Chapter 2

Historical Background

The sensation of hot and cold; that is temperature, and the use of fire in cooking (heat) predates written history. The scientific understanding of temperature and heat is surprisingly recent. Emilio Segrè has written a wonderful account of the evolution of our understanding of heat, temperature and other aspects of the foundation of modern physics.

The concepts of temperature, heat, work and entropy grew gradually from the time of Galileo and culminated with the formulation of classical thermodynamics by Rudolf Clausius (1822–1888). At this stage there were two *principles of thermodynamics* formulated by Clausius:

I The energy of the universe is constant

II The entropy of the universe tends to a maximum

Clausius introduce the word *entropy* in 1865,from the Greek τροπὴ (*Verwandlung*; transformation). Entropy, a fundamental concept in thermodynamics, is involved in both principles. Now entropy is most readily understood in terms of statistical physics. In classical physics there is no reference point for entropy. Walther Hermann Nernst (1864–1941) formulated the third principle of thermodynamics:

III The entropy vanishes for a system at zero absolute temperature,

a conclusion Nernst reached by experimental investigations of vapors at high temperature, solids at low temperature and galvanic cells. He received the 1920 Nobel Prize for Chemistry for his discovery

Classical mechanics is reversible. For particles that interact with forces that may be derived from a potential the classical paths are strictly time-reversible, and the total energy of the system is conserved. This situation is not changed by quantum mechanics. The deep problem is therefore to understand how entropy can always increase as equilibrium is approached.

Interestingly the principle of energy conservation, i. e. the so called *first law of thermodynamics* I, was established *after* Carnot's discovery of the second law of thermodynamics.

Entropy is a concept of great practical importance, but it is a concept that is difficult to understand. The following quote from a nice little book: *Steam Engine Principles*by Calvert (1991) illustrates the difficulty:

> A forth property is entropy.[1]This has no physical existence and it has an arbitrary datum, but its changes can be precisely calculated. Such changes are listed in tables and plotted on charts. It is unprofitable to find an analogue for entropy. There is none. It must be accepted and used. It is as much a handicap to try to understand what is going on in our steam engine without recourse to entropy tables or charts as it is for a coastwise navigator to reject the tide tables. In both cases the tables are prepared by the learned and they are intended to be used without question by the practical man.

Of course, there are objections to Clausius' formulation of the first law: The energy universe is not conserved unless we take into account the equivalence of mass and energy as stated in Einstein's famous formula $E = mc^2$. However, in classical systems that are 'closed' energy is indeed conserved.

There are equivalent formulations of these two 'laws' of thermodynamics that give more insight.

2.1 Temperature

Galileo (1564–1642) around 1592 built a fluid based *thermoscope* which could be used to show changes in temperature. However the thermoscope had no fixed point and temperatures form different thermoscopes could not be compared. Disciples of Galileo and others in the *Academia del Cimento* (in Florence) made systematic studies using a *thermometer*. Isaac Newton (1642–1727) proposed a temperature scale in which the freezing point of water was set to zero, and the temperature of the human body to twelve.

Daniel Gabriel Fahrenheit (1686–1736) was born in Danzig but studied in Holland and worked in Amsterdam. Fahrenheit proposed a temperature scale still used today. He used alcohol thermometers and was the first to make a mercury thermometer. Fahrenheit used three *fixed points*: (1): 0 degrees for a mixture of ice, water and salt; (2): 32 degrees for a mixture of ice and water; (3): 96 degrees at the body temperature of a 'healthy' person in the mouth or under the arm. This temperature scale has the advantage that most temperatures experienced in daily life are all positive.

Fahrenheit was the first to observe and describe super-cooling of water in small capillaries. He used rain-water in a 1 inch bulb that he evacuated in the manner used by him to make thermometers. Fahrenheit found that the water was liquid at 15 degrees, i. e. -9.4 °C, and solidified immediately if he broke the tip of the capillary attached to the bulb to admit air.

We now mostly use the Celsius temperature scale. Anders Celsius (1701–1744) was a professor of astronomy at the university of Uppsala. In 1742 he described his thermometer to the Swedish Academy of Sciences. The thermometer had the melting point of snow as one fixed point at 100 °C and the boiling point of water was set to 0 °C. This was later inverted so that 0° is the melting point of snow, and 100° as the other, for the centigrade scale. The name was changed in 1948 to honor Celsius.

---

[1] the others are: pressure, temperature, and density

William Thomson (later Lord Kelvin) introduce the absolute temperature scale. (More later)

2.2 Steam Engines

The steam engine was developed and played an important role in the industrial revolution, starting in England in the mid-18th century. The name we normally associate with the steam engine is James Watt (1736–1819) who introduced the *condenser*, patented in 1768: "A New Invented Method of Lessening the Consumption of Steam and Fuel in Fire Engines."[2] Watts invention greatly increased the efficiency[3] of the steam engine patented in 1705 by Thomas Newcomen (1663–1729) used to drive pumps in coal mines (see Fig. 2.1.) The steam engine was further developed in the coming years by many engineers without a deeper understanding of the efficiency of the engine. Watt's invention alone significantly reduced the amount of coal needed to perform a given amount of work, and thus made steam engines practical. There were dangers, of course. For instance, in 1877 there were in Germany 20 explosions of steam boilers injuring 58 persons. In 1890 there were 14 such explosions (of a vastly larger number of boilers) injuring 18 persons. Stationary steam engines were used in industry, for railroads, in agriculture and later for electric power generation. The table shows that there was a dramatic increase in the use of steam engines only 120 years ago—now they are almost extinct. Steam turbines have replaced steam engines completely for electric power generation.

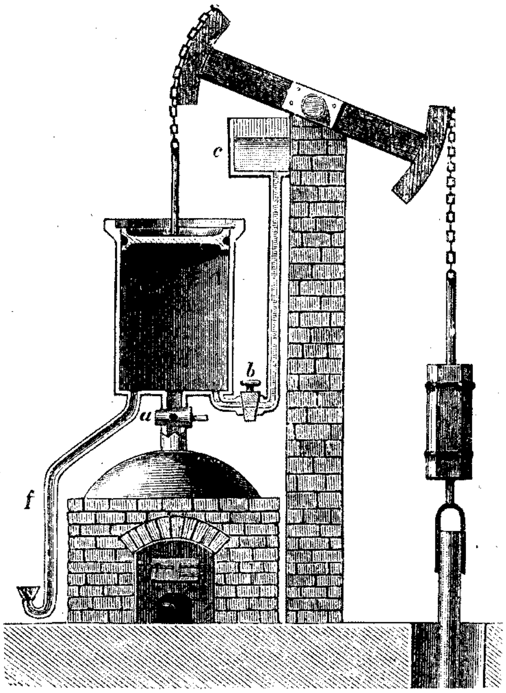

Figure 2.1: The Newcomen steam engine worked in a simple fashion. Steam generated in the boiler is admitted to the cylinder by opening the valve (a) and the weight of the other end of the arm pull the piston upward. The cylinder volume is filled and heated by the steam at atmospheric pressure. By closing the valve (a) and opening the valve (b) cold water is introduced into the cylinder. The steam condenses and cools thereby reducing the pressure in the cylinder so that the piston moves down pushed by the atmospheric pressure. The excess water runs out of the tube (f). When the cylinder has cooled down the cycle can start again. The maximum pressure difference that can be obtained is then given by the length of the tube (f) and the temperature of the surroundings.

---

[2] 'Watt, James' Britannica Online. <http://www.eb.com:180/cgi-bin/g?DocF=micro/632/91.html>

[3] from 1% to 2%

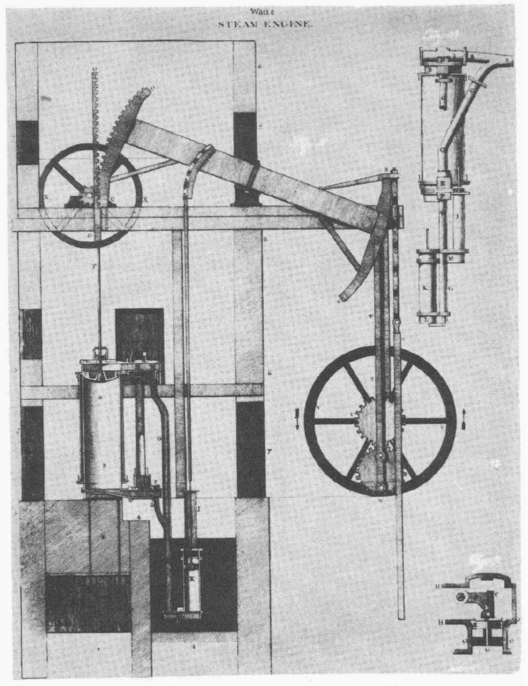

Figure 2.2: The steam was generated at atmospheric pressure and introduced into the cylinder just as in the Newcomen engine. However, in contrast with the Newcomen engine the steam was condensed in the small cylinder (K) immersed in cold water. The machine shown was *double-acting*,i. e. it worked both on the down- and the up-stroke. various arrangements opened and closed valves automatically. The water in the condenser was pumped out. The central improvement was that there was no need to cool the cylinder between strokes.

	1879		1879	
Type of Steam engine	Number	Horse power	Number	Horse power
Free standing	29 895	887 780	45 192	1 508 195
Mobile	5 442	47 104	11 916	111 070
Ship	623	50 309	1 674	154 189

Table 2.1: Steam engines in Prussia. Globally there were about 1.9 million steam engines in the world around 1890.

2.3 Sadi Carnot

Thermodynamics as a field of science started with Sadi Carnot (1796–1831). He published in 1824 remarkable book*Réflexions sur la Puissance Motrice du Feu, et sur les Machines Propres a développer cette Puissance*—his only published scientific work. But hardly anyone bought the book, and it became very difficult to obtain only a few years later. It was Émile Clapeyron (1799–1864), who understood Carnot's work and expressed the results in a more mathematical form in the paper published in 1834. Fourteen years later Clayperon's work was expanded by William Thomson[4] (1824–1907). Thomson studied Carnot's paper in 1847 and realized that the Carnot cycle could be used to define an *absolute temperature* scale. However, with Carnot, he as late as in 1848 thought that heat was conserved, which is incorrect. This view was, however, changed by James Prescott Joule (1818–1889), and together they established the correct interpretation and thereby the principle of *conservation of energy*.

---

[4] Thomson was knighted in 1866, raised to the peerage in 1892 as Baron Kelvin of Largs.

The beginning of Sadi Carnot's book gives a feel for the technological optimism at the beginning of the 19th century, and the motivation of his work:

> EVERY one knows that heat can produce motion. That it possesses vast motive-power no one can doubt, in these days when the steam-engine is everywhere so well known.
>
> To heat also are due the vast movements which take place on the earth. It causes the agitations of the atmosphere, the ascension of clouds, the fall of rain and of meteors,[5] the currents of water which channel the surface of the globe, and of which man has thus far employed but a small portion. Even earthquakes and volcanic eruptions are the result of heat.
>
> From this immense reservoir we may draw the moving force necessary for our purposes. Nature, in providing us with combustibles on all sides, has given us the power to produce, at all times and in all places, heat and the impelling power which is the result of it. To develop this power, to appropriate it to our uses, is the object of heat-engines.
>
> The study of these engines is of the greatest interest, their importance is enormous, their use is continually increasing, and they seem destined to produce a great revolution in the civilized world.
>
> Already the steam-engine works our mines, impels our ships, excavates our ports and our rivers, forges iron, fashions wood, grinds grains, spins and weaves our cloths, transports the heaviest burdens, etc. It appears that it must some day serve as a universal motor, and be substituted for animal power, waterfalls, and air currents.

After a discussion of the importance of steam-engines for England, he notes:

> The discovery of the steam-engine owed its birth, like most human inventions, to rude attempts which have been attributed to different persons, while the real author is not certainly known. It is, however, less in the first attempts that the principal discovery consists, than in the successive improvements which have brought steam-engines to the conditions in which we find them today. There is almost as great a distance between the first apparatus in which the expansive force of steam was displayed and the existing machine, as between the first raft that man ever made and the modern vessel.

Here Carnot expresses, in an amicable fashion, the unreasonable effectiveness of engineering. The path for the original invention to the final product of generations of engineers, through a never ending series of improvements leads to qualitatively new products.

After this motivation of his work he approaches the more general *scientific* question:

> Notwithstanding the work of all kinds done by steam-engines, notwithstanding the satisfactory condition to which they have been brought today, their theory is very little understood, and the attempts to improve them are still directed almost by chance.

---

[5] This is a misunderstanding, the movement of meteors has nothing to do with heat. When they enter the atmosphere, friction from the air cause them to heat and glow.

> The question has often been raised whether the motive power of heat[6] is unbounded, whether the possible improvements in steam engines have an assignable limit—a limit which the nature of things will not allow to be passed by any means whatever; or whether, on the contrary, these improvements may be carried on indefinitely. We have long sought, and are seeking today, to ascertain whether there are in existence agents preferable to the vapor of water for developing the motive power of heat; whether atmospheric air, for example, would not present in this respect great advantages. We propose now to submit these questions to a deliberate examination.

Then he really becomes precise and defines the scientific question, which he the actually answers:

> The phenomenon of the production of motion by heat has not been considered from a sufficiently general point of view. We have considered it only in machines the nature and mode of action of which have not allowed us to take in the whole extent of application of which it is susceptible. In such machines the phenomenon is, in a way, incomplete. It becomes difficult to recognize its principles and study its laws.
>
> In order to consider in the most general way the principle of the production of motion by heat, it must be considered independently of any mechanism or any particular agent. It is necessary to establish principles applicable not only to steam-engines[7] but to all imaginable heat-engines, whatever the working substance and whatever the method by which it is operated.

Carnot goes on to describe the workings of the steam-engine, and concludes:

> The production of motive power is then due in steam-engines not to an actual consumption of caloric, but to its transportation from a warm body to a cold body, that is, to its re-establishment of equilibrium an equilibrium considered as destroyed by any cause whatever, by chemical action such as combustion, or by any other. We shall see shortly that this principle is applicable to any machine set in motion by heat.
>
> ...
>
> We have shown that in steam-engines the motive-power is due to a re-establishment of equilibrium in the caloric; this takes place not only for steam-engines, but also for every heat-engine—that is, for every machine of which caloric is the motor. Heat can evidently be a cause of motion only by virtue of the changes of volume or of form which it produces in bodies.

---

[6] We use here the expression motive power to express the useful effect that a motor is capable of producing. This effect can always be likened to the elevation of a weight to a certain height. It has, as we know, as a measure, the product of the weight multiplied by the height to which it is raised. [*Here, in a footnote, Carnot gives a clear definition of what we today call free energy, even though his concept of heat (or caloric) was inadequate*]

[7] We distinguish here the steam-engine from the heat-engine in general. The latter may make use of any agent whatever, of the vapor of water or of any other to develop the motive power of heat.

Here, Carnot has unclear concepts about the role of heat, or caloric. in a footnote (see below) he evades the question. However, he is clear in the sense that a temperature difference is required if work is to be generated. In that process there is caloric or heat flowing from the hot to the cold, trying to re-establish thermal equilibrium.

Then Carnot rephrases the problem he is considering:

> It is natural to ask here this curious and important question: Is the motive power of heat invariable in quantity, or does it vary with the agent employed to realize it as the intermediary substance, selected as the subject of action of the heat?
>
> It is clear that this question can be asked only in regard to a given quantity of caloric,[8] the difference of the temperatures also being given. We take, for example, one body A kept at a temperature of 100° and another body B kept at a temperature of 0°, and ask what quantity of motive power can be produced by the passage of a given portion of caloric (for example, as much as is necessary to melt a kilogram of ice) from the first of these bodies to the second. We inquire whether this quantity of motive power is necessarily limited, whether it varies with the substance employed to realize it, whether the vapor of water offers in this respect more or less advantage than the vapor of alcohol, of mercury, a permanent gas, or any other substance. We will try to answer these questions, availing ourselves of ideas already established.
>
> We have already remarked upon this self-evident fact, or fact which at least appears evident as soon as we reflect on the changes of volume occasioned by heat: *wherever there exists a difference of temperature, motive power can be produced*.

Here he announced the basic principle: A temperature difference is required in order to generate work. Then he presents the first sketch of the cycle:

> Imagine two bodies A and B, kept each at a constant temperature, that of A being higher than that of B. These two bodies, to which we can give or from which we can remove the heat without causing their temperatures to vary, exercise the functions of two unlimited reservoirs of caloric. We will call the first the furnace and the second the refrigerator.
>
> If we wish to produce motive power by carrying a certain quantity of heat from the body A to the body B we shall proceed as follows:
>
> (1) To borrow caloric from the body A to make steam with it—that is, to make this body fulfill the function of a furnace, or rather of the metal composing the boiler in ordinary engines—we here assume that the steam is produced at the same temperature as the body A.
>
> (2) The steam having been received in a space capable of expansion, such as a cylinder furnished with a piston, to increase the volume of this space, and

---

[8] It is considered unnecessary to explain here what is quantity of caloric or quantity of heat (for we employ these two expressions indifferently), or to describe how we measure these quantities by the calorimeter. Nor will we explain what is meant by latent heat, degree of temperature, specific heat, etc. The reader should be familiarized with these terms through the study of the elementary treatises of physics or of chemistry.

> consequently also that of the steam. Thus rarefied, the temperature will fall spontaneously, as occurs with all elastic fluids; admit that the rarefaction may be continued to the point where the temperature becomes precisely that of the body B.
>
> (3) To condense the steam by putting it in contact with the body B, and at the same time exerting on it a constant pressure until it is entirely liquefied. The body B fills here the place of the injectionwater in ordinary engines, with this difference, that it condenses the vapor without mingling with it, and without changing its own temperature.

Here, Carnot introduced several of the basic concepts of Thermodynamics: *heat reservoir* and *thermodynamics process*. In (1) and (2) he describes an *isothermal process*, whereas in (2) he describes and *isentropic process*. Reversibility *reversibility* is the next concept:

> The operations which we have just described might have been performed in an inverse direction and order. There is nothing to prevent forming vapor with the caloric of the body B, and at the temperature of that body, compressing it in such a way as to make it acquire the temperature of the body A, finally condensing it by contact with this latter body, and continuing the compression to complete liquefaction.

This then completes the Carnot cycle. However, Carnot continues in a fashion that is difficult to understand, and in contradiction with what he writes later. First the continuation:

> By our first operations there would have been at the same time production of motive power and transfer of caloric from the body A to the body B. By the inverse operations there is at the same time expenditure of motive power and return of caloric from the body B to the body A. But if we have acted in each case on the same quantity of vapor, if there is produced no loss either of motive power or caloric, the quantity of motive power produced in the first place will be equal to that which would have been expended in the second, and the quantity of caloric passed in the first case from the body A to the body B would be equal to the quantity which passes back again in the second from the body B to the body A; so that an indefinite number of alternative operations of this sort could be carried on without in the end having either produced motive power or transferred caloric from one body to the other.

Here, by the assumption that caloric is conserved, he also gets no work done by his cycle! Nevertheless he an-ounces his second principle:

> Now if there existed any means of using heat preferable to those which we have employed, that is, if it were possible by any method whatever to make the caloric produce a quantity of motive power greater than we have made it produce by our first series of operations, it would suffice to divert a portion of this power in order by the method just indicated to make the caloric of the body B return to the body A from the refrigerator to the furnace, to restore the initial conditions, and thus to be ready to commence again an operation precisely similar to the former, and so

> on: this would be not only perpetual motion, but an unlimited creation of motive power without consumption either of caloric or of any other agent whatever. Such a creation is entirely contrary to ideas now accepted, to the laws of mechanics and of sound physics. It is inadmissible. We should then conclude that *the maximum of motive power resulting from the employment of steam is also the maximum of motive power realizable by any means whatever*.
>
> . . .
>
> Since every re-establishment of equilibrium in the caloric may be the cause of the production of motive power, every re-establishment of equilibrium which shall be accomplished without production of this power should be considered as an actual loss. Now, very little reflection would show that all change of temperature which is not due to a change of volume of the bodies can be only a useless re-establishment of equilibrium in the caloric. *The necessary condition of the maximum is, then, that in the bodies employed to realize the motive power of heat there should not occur any change of temperature which may not be due to a change of volume*.

Here he has two more fundamental concepts: There is am maximum work attainable that is independent of what type of heat-engine is employed. Secondly he notes that any heat conduction results in a loss of work as compared with the work produced by an ideal heat-engine. The problem is to understands how he reaches these correct conclusions, based on a cycle that produces neither work nor a transport of caloric (heat) from the hot to the cold reservoir.

Let us follow his argument for the second demonstration: First the basic assumption:

> When a gaseous fluid is rapidly compressed its temperature rises. It falls, on the contrary, when it is rapidly dilated. This is one of the facts best demonstrated by experiment. We will take it for the basis of our demonstration.

Then he goes on to describe in detail what we now call the Carnot cycle:

> This preliminary idea being established, let us imagine an elastic fluid, atmospheric air for example, shut up in a cylindrical vessel, *abcd* (Fig. 1), provided with a movable diaphragm or piston, *cd*. Let there be also two bodies, A and B, kept each at a constant temperature, that of A being higher than that of B. Let us picture to ourselves now the series of operations which are to be described, Fig. 2.3

This is the complete *Carnot cycle*. Carnot continues to explain that work is done:

> In these various operations the piston is subject to an effort of greater or less magnitude, exerted by the air enclosed in the cylinder; the elastic force of this air varies as much by reason of the changes in volume as of changes of temperature. But it should be remarked that with equal volumes, that is, for the similar positions of the piston, the temperature is higher during the movements of dilatation than during the movements of compression. During the former the elastic force of the air is found to be greater, and consequently the quantity of motive power produced by the movements of dilatation is more considerable than that consumed to produce the movements of

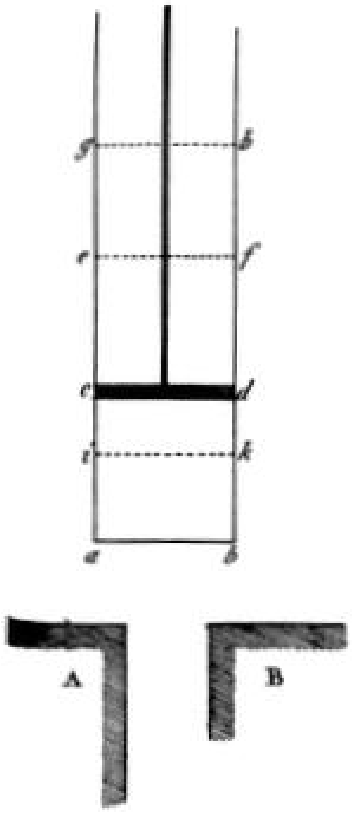

Figure 2.3: **(1)** Contact of the body A with the air enclosed in the space *abcd* or with the wall of this space—a wall that we will suppose to transmit the caloric readily. The air becomes by such contact of the same temperature as the body A; *cd* is the actual position of the piston.
(2) The piston gradually rises and takes the position *ef*. The body A is all the time in contact with the air, which is thus kept at a constant temperature during the rarefaction. The body A furnishes the caloric necessary to keep the temperature constant.
(3) The body A is removed, and the air is then no longer in contact with any body capable of furnishing it with caloric. The piston meanwhile continues to move, and passes from the position *ef* to the position *gh*. The air is rarefied without receiving caloric, and its temperature falls. Let us imagine that it falls thus till it becomes equal to that of the body B; at this instant the piston stops, remaining at the position *gh*.
(4) The air is placed in contact with the body B; it is compressed by the return of the piston as it is moved from the position *gh* to the position *cd* . This air remains, however, at a constant temperature because of its contact with the body B, to which it yields its caloric.
(5) The body B is removed, and the compression of the air is continued, which being then isolated, its temperature rises. The compression is continued till the air acquires the temperature of the body A. The piston passes during this time from the position *cd* to the position *ik*.**(6)** The air is again placed in contact with the body A. The piston returns from the position *ik* to the position *ef*; the temperature remains unchanged.
(7) The step described under number (3) is renewed, then successively the steps (4), (5), (6), (3), (4), (5), (6), (3), (4), (5); and so on.

> compression. Thus we should obtain an excess of motive power—an excess which we could employ for any purpose whatever. The air, then, has served as a heat-engine; we have, in fact, employed it in the most advantageous manner possible, for no useless re-establishment of equilibrium has been effected in the caloric.

Then he points out that the cycle may be run in reverse and thereby by the use of mechanical work take heat from the lower temperature of B to the higher temperature of the Reservoir A, and he continues:

> The result of these first operations has been the production of a certain quantity of motive power and the removal of caloric from the body A to the body B. The result of the inverse operations is the consumption of the motive power produced and the return of the caloric from the body B to the body A; so that these two series of operations annul each other, after a fashion, one neutralizing the other.
>
> The impossibility of making the caloric produce a greater quantity of motive power than that which we obtained from it by our first series of operations, is now easily proved. It is demonstrated by reasoning very similar to that employed at page 8; the reasoning will here be even more exact. The air which we have used to develop the motive power is restored at the end of each cycle of operations exactly to the state in which it was at first found, while, as we have already remarked, this would not be precisely the case with the vapor of water.[9]
>
> We have chosen atmospheric air as the instrument which should develop the motive power of heat, but it is evident that the reasoning would have been the same for all other gaseous substances, and even for all other bodies susceptible of change of temperature through successive contractions and dilatations, which comprehends all natural substances, or at least all those which are adapted to realize the motive power of heat. Thus we are led to establish this general proposition:
>
> *The motive power of heat is independent of the agents employed to realize it; its quantity is fixed solely by the temperatures of the bodies between which is ejected, finally, the transfer of the caloric.*
>
> We must understand here that each of the methods of developing motive power attains the perfection of which it is susceptible. This condition is found to be fulfilled if, as we remarked above, there is produced in the body no other change of temperature than that due to change of volume, or, what is the same thing in other words, if there is no contact between bodies of sensibly different temperatures.

---

[9] We tacitly assume in our demonstration, that when a body has experienced any changes, and when after a certain number of transformations it returns to precisely its original state, that is, to that state considered in respect to density, to temperature, to mode of aggregation-let us suppose, I say, that this body is found to contain the same quantity of heat that it contained at first, or else that the quantities of heat absorbed or set free in these different transformations are exactly compensated. This fact has never been called in question. It was first admitted without reflection, and verified afterwards in many cases by experiments with the calorimeter. To deny it would be to overthrow the whole theory of heat to which it serves as a basis. For the rest, we may say in passing, the main principles on which the theory of heat rests require the most careful examination. Many experimental facts appear almost inexplicable in the present state of this theory.

Here we are the final conclusion. The amount of work that may be extracted by a heat-engine, of any design whatever, depends only on the temperatures of the hot and the cold reservoir, and on the amount of heat transported. This conclusion is correct. However, the view that the *caloric* is conserved is clearly incorrect. The point is that the conservation of energy was only established much later. Carnot was clearly aware that his idea of heat was not clear as may be seen from his remarks:

> According to established principles at the present time, we can compare with sufficient accuracy the motive power of heat to that of a waterfall. Each has a maximum that we cannot exceed, whatever may be, on the one hand, the machine which is acted upon by the water, and whatever, on the other hand, the substance acted upon by the heat. The motive power of a waterfall depends on its height and on the quantity of the liquid; the motive power of heat depends also on the quantity of caloric used, and on what may be termed, on what in fact we will call, *the height of its fall*,[10] that is to say, the difference of temperature of the bodies between which the exchange of caloric is made. In the waterfall the motive power is exactly proportional to the difference of level between the higher and lower reservoirs. In the fall of caloric the motive power undoubtedly increases with the difference of temperature between the warm and the cold bodies; but we do not know whether it is proportional to this difference. We do not know, for example, whether the fall of caloric from 100 to 50 degrees furnishes more or less motive power than the fall of this same caloric from 50 to zero. It is a question which we propose to examine hereafter.

Carnot thought that caloric was a conserved quantity—an attitude consistent with the experimental accuracy of his day. The correct interpretation was first given by Rudolf Clausius in 1850,with the formulation:

> *dass in allen Fällen, wo durch Wärme Arbeit entstehe, eine der erzeugten Arbeit proportionale Wärmemenge verbraucht werde, und dass umgekehrt duch Verbrauch einer ebenso grossen Arbeit dieselbe Wärmemenge erzeugt werden könne.*

2.4 The Maxwell's Distribution

Maxwell[11] read a paper at the meeting of the British Association of Aberdeen in September 1859. It was published in the *Philosophical Magazine* in 1860, and it contains the following

---

[10] The matter here dealt with being entirely new, we are obliged to employ expressions not in use as yet, and which perhaps are less clear than is desirable.

[11] Maxwell, James Clerk; born June 13, 1831, Edinburgh, Scotland; died November 5, 1879, Cambridge, Cambridgeshire, England; Scottish physicist best known for his formulation of electromagnetic theory. He is regarded by most modern physicists as the scientist of the 19th century who had the greatest influence on 20th-century physics, and he is ranked with Sir Isaac Newton and Albert Einstein for the fundamental nature of his contributions. In 1931, on the 100th anniversary of Maxwell's birth, Einstein described the change in the conception of reality in physics that resulted from Maxwell's work as ”the most profound and the most fruitful that physics has experienced since the time of Newton.”

heuristic argument on the velocity distribution.[12]

> If a great many equal spherical particles were in motion in a perfectly elastic vessel, collisions would take place among the particles, and their velocities would be altered at every collision; so that after a certain time the *vis viva* will be divided among the particles according to some regular law, the average number of particles whose velocity lies between certain limits being ascertainable though the velocity of each particle changes at every collision.
>
> Prop. IV. To find the average number of particles whose velocities lie between given limits, after a great number of collisions among a great number of equal particles.
>
> Let N be the whole number of particles. Let x, y, z be the components of the velocity of each particle in three rectangular directions, and let the number of particles for which x lies between x and $x+dx$, be $Nf(x)dx$, where $f(x)$ is a function of x to be determined.
>
> The number of particles for which y lies between y and $y+dy$ will be $Nf(y)dy$; and the number for which z lies between z and $z+dz$ will be $Nf(z)dz$ where f always stands for the same function.
>
> Now the existence of the velocity x does not in any way affect that of the velocities y or z, since these are all at right angles to each other and independent, so that the number of particles whose velocity lies between x and $x+dx$, and also between y and $y+dy$, and also between z and $z+dz$, is
>
> $$Nf(x)f(y)f(z)dxdydz\ .$$
>
> If we suppose the N particles to start from the origin at the same instant, then this will be the number in the element of volume $(dxdydz)$ after unit of time, and the number referred to unit of volume will be
>
> $$Nf(x)f(y)f(z)\ .$$
>
> But the directions of the coordinates are perfectly arbitrary, and therefore this number must depend on the distance from the origin alone, that is
>
> $$f(x)f(y)f(z) = \phi(x^2+y^2+z^2)\ .$$
>
> Solving this functional equation, we find
>
> $$f(x) = Ce^{Ax^2}\ , \qquad \phi(r^2) = C^3 e^{Ar^2}\ .$$
>
> If we make A positive, the number of particles will increase with the velocity, and we should find the whole number of particles infinite. We therefore make A negative and equal to $-\frac{1}{\alpha^2}$, so that the number between x and $x+dx$ is

---

[12] Appendix 10 from Emilio Segrè's book Maxwell's Distribution of Velocities of Molecules in His Own Words

$$NCe^{-x^2/\alpha^2}dx$$

Integrating from $x = -\infty$ to $x = +\infty$, we find the whole number of particles,

$$NC\sqrt{\pi}\alpha = N \text{ , therefore } C = \frac{1}{\alpha\sqrt{\pi}} \text{ ,}$$

$f(x)$ is therefore

$$\frac{1}{\alpha\sqrt{\pi}}e^{-x^2/\alpha^2}$$

Whence we may draw the following conclusions:—

1st. The number of particles whose velocity, resolved in a certain direction, lies between x and $x + dx$ is

$$N\frac{1}{\alpha\sqrt{\pi}}e^{-x^2/\alpha^2}dx$$

2nd. The number whose actual velocity lies between v and $v + dv$ is

$$N\frac{1}{\alpha^3\sqrt{\pi}}v^2e^{-v^2/\alpha^2}dv$$

3rd. To find the mean value of v, add the velocities of all the particles together and divide by the number of particles; the result is

$$\text{mean velocity} = \frac{2\alpha}{\sqrt{\pi}}$$

4th. To find the mean value of v^2, add all the values together and divide by N,

$$\text{mean value of } v^2 = \frac{3}{2}\alpha^2 \text{ .}$$

This is greater than the square of the mean velocity, as it ought to be.

Chapter 3

Basic principles of statistical mechanics

3.1 Microscopic states

Classical phase space: *Classical phase space* is $2f$-dimensional space

$$(q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f)\,.$$

Here q_i are *positional* coordinates, p_i are *momenta*, $f = 3N$ for n independent particles in 3-dimensional space. Each point corresponds to a *microscopic state* of the system. The simplest Hamiltonian of such a system is

$$\mathcal{H} = \sum_j^f \frac{p_j^2}{2m} + E(q_1, q_2, \ldots, q_f)\,. \tag{3.1}$$

Motion of the system is determined by the canonical equation of motion,

$$\dot{p}_j = -\frac{\partial \mathcal{H}}{\partial q_j}\,, \quad \dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}\,, \quad (j = 1, 2, \ldots, f)\,. \tag{3.2}$$

The motion of the *phase point* $\mathbf{P}_t$ defines the state at time t. The trajectory of the phase space is called the *phase trajectory*. For a given initial conditions at some $t - t_0$, $q_0 \equiv \{q_{j0}\},\ p_0 \equiv \{p_{j0}\}$, the phase trajectory is given by the functions $q_i(t, q_0, p_0),\ p_i(t, q_0, p_0)$.

For a conservative system energy is conserved,

$$\mathcal{H}(q, p) = E\,. \tag{3.3}$$

Thus the orbit belongs to the surface of constant energy.

Exercise: Analyse phase space for a one-dimensional harmonic oscillator and find the dependence of its energy on the amplitude..

Quantum states: *Quantum states* ψ_l are defined by the Schrödinder equation ,

$$\mathcal{H}\psi_l = E_l \psi_l \quad (l = 1, 2, \ldots)\,. \tag{3.4}$$

They are determined by the *quantum numbers*. We will discuss quantum systems later.

3.2 Statistical treatment

Fundamental assumption: Let A be a physical quantity dependent on the microscopic state. In classical mechanics,

$$A(q,p) = A(\mathbf{P})\,.$$

In quantum mechanics,

$$A_l = \int \psi_l^* A \psi_l \,(dq) \equiv \langle l|A|l\rangle\,.$$

Here $(dq) \equiv \prod_{j=1}^{f} dq_j$. The observed value (in a macroscopic sense), A_{obs} *must be a certain average* of microscopic A,

$$A_{\text{obs}} = \bar{A}\,.$$

Realization probability of a microscopic state: Let $\mathcal{M}$ be the set of microscopic states which can be realized under a certain microscopic condition. The *realization probability* that one of the microscopic states belongs to the element $\Delta\Gamma$ of the phase space is defined as

$$\int_{\Delta\Gamma} f(\mathbf{P})\,d\Gamma\,, \quad (\Delta\Gamma \in \mathcal{M})\,.$$

The probability that the quantum state l is realized is defined as

$$f(l)\,, \quad (l \in \mathcal{M})\,.$$

f is usually called the *distribution functions*.

The observed quantity is given by the formulas

$$\begin{aligned} A_{\text{obs}} = \bar{A} &= \int_{\mathcal{M}} A(\mathbf{P})\, f(\mathbf{P})\,d\Gamma\,, \\ \bar{A} &= \sum_{\mathcal{M}} A_l\, f(l)\,. \end{aligned} \tag{3.5}$$

Statistical ensembles: great number of systems each of which has the same structure as the system under consideration. It is assumed that

$$A_{\text{obs}} = \text{ensemble average of } A = \bar{A}\,.$$

Example – ideal gas: The phase space consists of spatial coordinates $\mathbf{r}$ and moments $\mathbf{p}$. In thermal equilibrium, at high temperature, and for a dilute gas, the *Maxwell distribution* reads:

$$f(\mathbf{p}) = (2\pi m k_B T)^{-1/2} \exp\left[-\frac{\mathbf{p}^2}{2mk_BT}\right]\,, \tag{3.6}$$

where T is the absolute temperatures, m the mass of a molecule, and k_B is the Boltzmann constant.

General properties of statistical distributions

Statistical independence: For two *macroscopic* and *non-interacting* subsystems,

$$f_{12}\,dp^{(12)}dq^{(12)} = f_1\,dp^{(1)}dq^{(1)}\;f_2\,dp^{(2)}dq^{(2)} \quad \to \quad f_{12} = f_1 f_2\,.$$

Statistical independence means

$$\overline{A_{12}} = \overline{A_1}\,\overline{A_2}\,,$$

where A is any physical quantity. Thus, for any additive quantity,

$$\bar{A} = \sum_j \bar{A}_j\,.$$

If one defines the fluctuation, $\Delta A = A - \bar{A}$, then

$$\langle(\Delta A)^2\rangle = \left\langle \left(\sum_j^N \Delta A_j\right)^2 \right\rangle = \sum_j^N \langle(\Delta A_j)^2\rangle\,.$$

As a result,

$$\frac{\langle(\Delta A)^2\rangle^{1/2}}{\bar{A}} \sim \frac{1}{\sqrt{N}} \ll 1\,.$$

Liouville theorem: Consider a set of instances, $t_1, t_2, \ldots,$. The states at that instances are represented by the points $\mathbf{P}_1, \mathbf{P}_2, \ldots,$. The *Liouville theorem* states that the distribution function $f(p,q)$ is *time-independent*. Indeed, in the absence of external forces, the state must be stationary, and for the phase-space points one has

$$\operatorname{div}(f\mathbf{v}) = 0 \quad \to \quad \sum_{i=1}^{f}\left[\frac{\partial}{\partial q_i}(f\dot{q}_i) + \frac{\partial}{\partial p_i}(f\dot{p}_i)\right] = 0\,.$$

Expanding derivatives we get

$$\sum_{i=1}^{f}\left[\dot{q}_i\frac{\partial f}{\partial q_i} + \dot{p}_i\frac{\partial f}{\partial p_i}\right] + f\sum_{i=1}^{f}\left[\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i}\right] = 0\,.$$

According to the Hamilton equation (3.2),

$$\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} = 0\,.$$

Thus

$$\frac{df}{dt} = \sum_{i=1}^{f}\left[\dot{q}_i\frac{\partial f}{\partial q_i} + \dot{p}_i\frac{\partial f}{\partial p_i}\right] = 0\,. \tag{3.7}$$

According to the Hamilton equation (3.2) the sum can be rewritten as

$$\sum_{i=1}^{f}\left[\dot{q}_i\frac{\partial f}{\partial q_i} + \dot{p}_i\frac{\partial f}{\partial p_i}\right] = \sum_{i=1}^{f}\left[\frac{\partial f}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i}\right] \equiv \{f, \mathcal{H}\}\,. \tag{3.8}$$

This combination is called the *Poisson bracket*.

Role of integrals of motion Since $\log f(p.q)$ must be (i) stationary, and (ii) additive for macroscopic subsystems of the system,

$$\log f_{12} = \log f_1 + \log f_2 ,$$

it depends only on additive integrals of motion, $E(p.q)$, total momentum $\mathbf{P}(p,q)$ and angular momentum $\mathbf{M}(p.q)$. The only additive function is the linear combination,

$$\log f(p,q) = \alpha + \beta E(p,q) + \vec{\gamma} \cdot \mathbf{P}(p,q) + \vec{\delta} \cdot \mathbf{M}(p,q) . \tag{3.9}$$

Here in the equilibrium the coefficients β, $\vec{\gamma}$ and $\vec{\delta}$ must be the same for all subsystems in the closed system. The coefficient α is defined from the normalization condition. Thus, the values of the integrals of motion completely define statistical properties of the closed system. In the following we assume that the system is enclosed in a rigid box, and $\mathbf{P} = 0$, $\mathbf{M} = 0$.

3.3 The principle of equal weight and the microcanonical ensemble

An *isolated system* after some time reaches the equilibrium. The energy of the system, E, is fixed with some allowance, δE. In a classical system the set $\mathcal{M}(E, \delta E)$ is the shell-like subspace of the phase space between the constant-energy surfaces for $\mathcal{H} = E$ and $\mathcal{H} = E + \delta E$.

For quantum mechanics, it is a set of quantum states having energies in the interval $E < E_l < E + \delta E$.

The principle of equal weight: In a thermal equilibrium state of an isolated system, each of the microscopic states belonging to $\mathcal{M}(E, E + \delta E)$ is realized with equal probability:

$$\begin{aligned} f(\mathbf{P}) &= \text{constant} = \left[\int_{E<\mathcal{H}<E+\delta E} d\Gamma \right]^{-1} , \qquad \mathbf{P} \in \mathcal{M}(E, \delta E) , \\ f(l) &= \text{constant} = \left[\sum_{E<E_l<E+\delta E} 1 \right]^{-1} , \qquad l \in \mathcal{M}(E, \delta E) . \end{aligned} \tag{3.10}$$

This distribution is called the *microcanonical one*, the corresponding ensemble being called the *microcanonical*.[1]

The microcanonical ensemble represents an isolated system which has reached thermal equilibrium.

---

[1] Some other additive integrals of motion besides the energy can be important.

Classical limit ($\delta E \to 0$): One can take a set $\sigma(E)$ on the surface of constant energy. Then All the phase trajectories are concentrated at the surface

$$\mathcal{H}(p,q) = E\,,$$

and the distribution function can be written as

$$f(\mathbf{P}) = \text{const} \cdot \delta[\mathcal{H}(p,q) - E]\,. \tag{3.11}$$

To determine the constant we have to calculate the integral

$$D(E) = \int d\Gamma\, \delta[\mathcal{H}(p,q) - E] \tag{3.12}$$

which is called *the density of states*, see below. It is very convenient to use the finite-width layer in the phase space and to transform variables from $\{p,q\}$ to the element of surface σ defined by the equality $\mathcal{H}(p,q) - E$, and to the coordinate n along the normal to this surface. Then the energy is dependent only on the coordinate n, and

$$dE = \frac{\partial \mathcal{H}}{\partial n}\, dn = |\nabla \mathcal{H}|\, dn\,.$$

Here

$$|\nabla \mathcal{H}| = \left[\sum_j \left\{ \left(\frac{\partial \mathcal{H}}{\partial p_j}\right)^2 + \left(\frac{\partial \mathcal{H}}{\partial q_j}\right)^2 \right\}\right]^{1/2}. \tag{3.13}$$

Consequently,

$$d\Gamma\, \delta[\mathcal{H}(p,q) - E] = \frac{d\sigma\, dn}{|\nabla \mathcal{H}|}\, \delta(n)\,.$$

Then one can perform integral over n and introduce the distribution over the constant energy surface,

$$f(\mathbf{P})\, d\sigma = \frac{1}{D(E)} \frac{d\sigma}{|\nabla \mathcal{H}|}\,, \quad D(E) = \left[\int_{\mathcal{H}=E} \frac{d\sigma}{|\nabla \mathcal{H}|}\right]. \tag{3.14}$$

An average is then given by the equation,

$$\bar{A} = \frac{1}{D(E)} \int_{\mathcal{H}=E} \frac{A(\mathbf{P})\, d\sigma}{|\nabla \mathcal{H}|}\,. \tag{3.15}$$

Ergodic theorem: Originally, A_{obs} is assumed to be *time average* of $A(\mathbf{p}_t)$. The ergodic theorem can be formulated as

$$A_{\text{time average}} = A_{\text{phase average}}\,.$$

Actually one can deduce the principle of equivalent weight from this theorem.

The finite allowance of the energy δE: In quantum mechanics, the energy of a system has uncertainty,

$$(\delta E)_{\text{qu}} \sim \hbar/t \,,$$

where t is the length of the observation. The statistical treatment is still possible only if the set $\mathcal{M}\left[E, \delta E \geq (\delta E)_{\text{qu}}\right]$ contains many quantum states.

If the system occupies only a single quantum state, statistics is not applicable.

3.4 The thermodynamic weight of a macroscopic state and entropy

Variables defining a macroscopic state: One can choose the energy E (with allowance δE), the numbers $N_A, N_B, \ldots$ of particles of various kinds, the volume V of the box which contains the system, and other parameters $x, \ldots$ which specify external forces. Thus the Hamiltonian depends on those parameters, $\mathcal{H} = \mathcal{H}[E(\delta E), N_A, N_B, \ldots, V, x, \ldots]$. In a classical system one can tend the allowance δE to 0. It is effectively true also for a quantum system provided the measuring time is long enough. Then the following definition do not depend on δE.

Thermodynamic weight: In quantum mechanics, the total number of quantum states, $W(E, \delta E, N, V, x)$ of the possible quantum states for the set of prescribed values

$$E(\delta E), N_A, N_B, \ldots, V, x, \ldots$$

is called *the thermodynamic weight* of that particular macroscopic state. Namely

$$W(E, \delta E, N, V, x) \equiv \sum_{E < E_l(N,V,x) < E + \delta E} 1 \,. \tag{3.16}$$

In classical mechanics it is defined as a limit of quantum statistical mechanics

$$W(E, \delta E, N, V, x) \equiv \int_{E < \mathcal{H}(N,V,x) < E + \delta E} \frac{d\Gamma}{h^{3(N_A + N_B + \ldots)} N_A! N_B! \ldots} \,. \tag{3.17}$$

Here h is the Planck constant.

Statistical definition of entropy:

$$S(E, N_A, N_B, \ldots, V, x, \ldots) = k_B \log W(E, \delta E, N, V, x) \,. \tag{3.18}$$

This relation is called the Boltzmann relation while the constant

$$k_B = 1.38 \times 10^{-23}\ \text{J/K} \tag{3.19}$$

is called the *Boltzmann constant*. To convince yourself that the statistical definition of entropy (3.18) is compatible with thermodynamics one has to check thermodynamic relations.[2]

---

[2] Note that allowance δE does not affect the value of the entropy

3.5 Number of states and density of states

Let us chose the energies to be non-negative,

$$0 \le E_1 \le E_2 \le \dots .$$

The number of states between 0 and E is called the number of states of the system:

$$\Gamma(E,N,V,x) \equiv \sum_{0<E_l\le E} 1 . \tag{3.20}$$

In classical mechanics,

$$\Gamma(E,N,V,x) \equiv \int_{0<\mathcal{H}<E} \frac{d\Gamma}{h^{3(N_A+N_B+\dots)} N_A! N_B! \dots} . \tag{3.21}$$

State density:

$$D(E,N,V,x) \equiv \frac{\partial}{\partial E} \Gamma(E,N,V,x) . \tag{3.22}$$

If allowance δE is small,

$$D(E,N,V)\, \delta E = W(E,\delta E,N,V,x) .$$

Correspondence of classical and quantum mechanics: It is important to keep in mind

(i) Uncertainty condition,

$$\Delta p \Delta q \sim h .$$

Because of that the classical phase volume $\Delta\Gamma$ corresponds to $\Delta\Gamma/h^f$ quantum states (f is the number of degrees of freedom;

(ii) *Indistinguishability of identical particles:* it is an important features of Bose and Fermi statistics. In the classical limit $h \to 0$ it leads to a factor $1/\prod_i N_i!$.

We shall come back to that point later discussing quantum statistics in more detail. Though the factor $1/N!$ *arises only from quantum statistics*, its necessity has been understood much earlier to make the entropy extensive quantity.

After definition of number of states

Another definition of entropy: Another definition of entropy can be formulated through the probability distribution. For a closed system, the distribution function can be written as a function of the energy, $f_l = f(E_l)$. Then the required number of states in the region $E, E+dE$ can be written as $D(E)\,dE$, while the energy probability distribution is

$$\mathcal{P}(E) = f(E) D(E)\, dE .$$

Since $\int \mathcal{P}(E)\,dE = 1$ and $\mathcal{P}(E)$ is centered near the average energy,

$$\mathcal{P}(\bar{E})\Delta E \approx f(\bar{E})D(\bar{E})\Delta E \approx f(\bar{E})W = 1\,.$$

Thus

$$S = -k\log f(\bar{E}) = -k\langle \log f(E_l)\rangle\,. \tag{3.23}$$

The distribution function f is constrained by the normalization condition

$$\sum_i f(E_i) = 1\,. \tag{3.24}$$

The latter inequality follows from *statistical independence*. Indeed, $\log f(E_l)$ must be an additive function of the energy (in general, also of other additive integrals of motion, **P** and **M**),

$$\log f(E_l) = \alpha + \beta E_l\,.$$

As a result

$$\log f(\bar{E}) = \langle \log f(E_l)\rangle = -\sum_l f_l \log f(E_l)\,.$$

This expression is written for *quantum* statistics. In the *classical* statistics one has to remember the normalization factor $[h^f \prod_j N_j!]^{-1}$ for the thermodynamic weight. Thus one has to replace

$$f_l \to \left\{ h^f \prod_j N_j! \right\} f(p,q)$$

in the argument of the logarithm to obtain

$$S = -\int d\Gamma\, f(p,q)\, \log\left[h^f \left(\prod_j N_j! \right) f(p,q) \right]. \tag{3.25}$$

The entropy, defined in this way, is additive.

Physical significance of entropy: Let us start with the microcanonical distribution,

$$\begin{aligned} d\mathcal{P} &= \text{const} \times \delta(E - E_0) \cdot \prod_a d\Gamma_a \\ &= \text{const} \times \delta(E - E_0) \cdot \prod_a (d\Gamma_a/dE_a)\, dE_a \\ &\quad \text{(after replacement } d\Gamma_a/dE_a \to \Delta\Gamma_a/\Delta E_a) \\ &= \text{const} \times \delta(E - E_0)\, e^S \prod_a dE_a/\Delta E_a\,. \end{aligned}$$

Here $S = \sum_a S_a,\ E = \sum_a E_a$.

Partial entropy $S(E_1, E_2, \ldots)$ must have maximum at $E_a = \bar{E}_a$. Since these values correspond to the *equilibrium state* the entropy of a closed system in a state of complete statistical equilibrium has its greatest possible value (for a given energy of the system).

The law of increase of entropy: If a closed system is at some instant in a non-equilibrium macroscopic state, the most probable consequence at later instants is a steady increase of the entropy of the system.

This is the *law of increase of entropy* or *the second law of thermodynamics*. This law has been formulated by R. Clausius (1865), its statistical explanation was given by L. Boltzmann (1870). It has a character of assumption which cannot be rigorously proved because the equations of mechanics are *time-reversible*.

If the entropy remains constant. the process is called *reversible*, if it increases the process is called *irreversible*.

3.6 Information

The definition of entropy in Eq. (3.23), is directly related to the definition of the information measure. As an example consider the *probability distribution* for the number of 'eyes' that come up when throwing a fair die. The probabilities of the six possible states $|1\rangle, |2\rangle, \dots, |6\rangle$ are f_i, with $i = 1, \dots, 6$. Small variations δf_i, lead to a variation in S as defined in (3.23). The maximal value of S may be found, taking into account the *constraint* (3.24), using a *Lagrange parameter* $k_B\lambda$:

$$\begin{aligned}
0 &= \delta(S - k_B\lambda\cdot 1) = \delta\left(-\sum_{i=1}^{6} k_B f_i \ln f_i - k_B\lambda \sum_{i=1}^{6} f_i\right) \\
&= -k_B \sum_{i=1}^{6}\left(\ln f_i + \frac{f_i}{f_i} + \lambda\right)\delta f_i \\
\Rightarrow &\quad f_i = \exp(-1-\lambda) \;\Rightarrow\; f_i = \frac{1}{6}\,. \qquad (3.26)
\end{aligned}$$

Here we have chosen λ so that f_i satisfies the constraint (3.24). We see that requiring the *information entropy* to have its maximum value leads to the result that $f_i = 1/6$, independent of i—as expected for fair dice.

When we throw two dice, we have many more possibilities. As before, the 'first' die has the possible states $\{|i\rangle\}$, with $i = 1, \dots, 6$. The other die has the states $\{|j\rangle\}$ with $j = 1, \dots, 6$, of course. Counting the eyes for the pair of dice we have 36 possible outcomes $\{|\alpha\rangle\}$, with $\alpha = 1, \dots, 36$ corresponding to the states $|11\rangle, |12\rangle, \dots |66\rangle$. With the argument developed above we will find that $f_\alpha = 1/36$. It is important to note, however, that the *statistical independence* of the two dice allows another approach.

The states $|\alpha\rangle = |i\rangle|j\rangle$ are in a product space of the two independent states $|i\rangle$ and $|j\rangle$, much in the same way that the position of a point in the plane $\mathbf{r} = (x,y)$, is given in terms of the independent coordinates x and y; or with our notation $|\mathbf{r}\rangle = |x\rangle|y\rangle$. The process of throwing the dice yields eyes that are statistically independent, and the probability, f_{ij}, that the first die has i eyes and the second has j eyes is the product of the probability f_i for the first die to show i eyes and f_j, which is the probability that the second die shows j eyes:

$$f_\alpha = f_{ij} = f_i f_j \quad \text{statistically independent} \qquad (3.27)$$

It follows that the entropy of the combined system may be written

$$\begin{aligned} S &= -k_B \sum_{ij} f_{ij} \ln f_{ij} = -k_B \sum_{ij} f_i f_j \ln f_i f_j \\ &= -k_B \sum_i f_i \ln f_i \sum_j f_j + -k_B \sum_j f_j \ln f_j \sum_i f_i \\ &= S_1 + S_2 \,, \end{aligned}$$

where we have used that probabilities are normalized: $\sum_i f_i = \sum_j f_j = 1$. We conclude that the (information) entropy of a system consisting of statistically independent parts is the sum of the entropies of the parts. Entropy is an extensive property, i.e. entropy is proportional to system size.

3.7 Normal systems in statistical thermodynamics

Asymptotic forms for the number of states and state density:

(i) When number of particles N, or the volume V tends to infinity, the number of states in normal systems approach the following asymptotic behavior:

$$\begin{aligned} \Gamma \sim \exp[N\psi(E/N)] \qquad &\text{or} \qquad \exp[V\psi(E/V)]\,, \\ \Gamma \sim \exp[N\psi(E/N, V/N)] \qquad &\text{or} \qquad \exp[V\psi(E/V, N/V)], \end{aligned} \tag{3.28}$$

and

$$\psi > 0\,, \quad \psi' > 0\,, \quad \psi'' > 0\,. \tag{3.29}$$

(ii) Therefore

$$\begin{aligned} D(E) &= \frac{\partial \Gamma}{\partial E} = \psi' \exp(N\psi) > 0\,, \\ \frac{\partial D(E)}{\partial E} &\sim \psi'^2 \exp(N\psi) > 0\,. \end{aligned} \tag{3.30}$$

Entropy of a normal thermodynamic system: For a normal system, the statistical entropy is

$$S = k \log[D(E)\Delta E] \approx k \log \Gamma(E) = kN\psi\,. \tag{3.31}$$

It is proportional to N, or to V.

Chapter 4

Thermodynamic quantities

4.1 Temperature

Consider 2 bodies in a thermal equilibrium which are in thermal contact with each other but are isolated form the surrounding, i.e. form a closed system. Since both the energy and entropy are additive, we have

$$E = E_1 + E_2 , \quad S = S(E_1) + S(E_2) .$$

In the equilibrium the entropy has a maximum with respect to the parameters of subsystems. Consequently, $dS/dE_1 = 0$. Form this equality we get

$$\begin{aligned} \frac{dS}{dE_1} &= \frac{dS_1}{dE_1} + \frac{dS_2}{dE_2}\frac{dE_2}{dE_1} \\ &= \frac{dS_1}{dE_1} - \frac{dS_2}{dE_2} = 0 , \end{aligned}$$

Here we have used the energy conservation, $E_2 = E - E_1$. As a result,

$$\frac{dS_1}{dE_1} = \frac{dS_2}{dE_2} .$$

The `absolute temperature` is defined as $T = dE/dS$, or, more rigorously,

$$T = (dE/dS)_V \tag{4.1}$$

which means that the volume V is kept constant. In the equilibrium, $T_1 = T_2$.

The previous argument may be generalized to a system consisting of many parts. At equilibrium one concludes that $\partial S/\partial E$ has the same value everywhere. Therefore one finds that absolute temperature T, defined as the change in the (equilibrium) entropy with energy is independent of position for a system in thermal equilibrium; it is also the same for two systems in equilibrium with each other, for example for a liquid in equilibrium with its vapor. In Eq. (4.1) we have explicitly written that the change is to take place at a fixed volume. Volume changes are discussed in the section 4.3.

Now let us discuss the case when the bodies and not in an equilibrium with each other, $T_1 \neq T_2$. During the equilibration the entropy $S = S_1 + S_2$ must increase,

$$\begin{aligned} \frac{dS}{dt} &= \frac{dS_1}{dt} + \frac{dS_2}{dt} \\ &= \frac{dS_1}{dE_1}\frac{dE_1}{dt} + \frac{dS_2}{dE_2}\frac{dE_2}{dt} > 0 \, . \end{aligned}$$

Since the energy is conserved,

$$\frac{dS}{dt} = \left(\frac{dS_1}{dE_1} - \frac{dS_2}{dE_2}\right)\frac{dE_1}{dt} = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\frac{dE_1}{dt} > 0 \, .$$

If $T_2 > T_1$, then $dE_1/dt > 0$, $dE_2/dt < 0$ - energy passes from bodies at higher temperature to bodies at lower temperature.

The temperature is measured in *degrees Kelvin*, the number of Joules per degree is called *the Boltzmann constant*, see Eq. (3.19). Below we omit the Boltzmann constant in all the formulas. Then the temperature is measured in the units of energy. To get custom formulas one has to replace $T \to k_B T$, $S \to S/k_B$.

4.2 Adiabatic process

Consider a *thermally isolated* system under *time-dependent* force λ, the classical Hamiltonian being $\mathcal{H}(p,q,t)$. Thermally isolated means that $\mathcal{H}$ is dependent on the coordinates and moments of the molecules belonging *only* to the body in question. If the force is purely mechanical (we ignore the feedback to the force source from the system under consideration), then one can apply the law of increase of entropy to the body in question.

Let us then assume that λ varies in time *sufficiently slow*[1] and expand dS/dt in powers of $d\lambda/dt$. Up to the lowest order,

$$\frac{dS}{dt} = A\left(\frac{d\lambda}{dt}\right)^2 \quad \to \quad \frac{dS}{d\lambda} = A\frac{d\lambda}{dt} \, .$$

Indeed, the 1st power is absent because entropy must increase with time. We obtain that at $d\lambda/dt \to 0$ the entropy tends to be λ-independent. Thus, the *adiabatic process* appears *reversible*.

Using adiabatic process one can calculate mean values. For example, let us define the thermodynamic energy as

$$E \equiv \overline{\mathcal{H}(p,q;\lambda)} \, .$$

Since $d\mathcal{H}(p,q;\lambda)/dt = \partial\mathcal{H}(p,q;\lambda)/\partial t$ (that follows from the Hamilton equation),

$$\frac{dE}{dt} = \overline{\frac{d\mathcal{H}(p,q;\lambda)}{dt}} = \overline{\frac{\partial\mathcal{H}(p,q;\lambda)}{d\lambda}}\,\frac{d\lambda}{dt} \, .$$

---

[1] comparing to the relaxation time of the system

On the other hand, E is a function of the entropy, S, and of external parameter, λ. Thus

$$\frac{dE}{dt} = \left(\frac{\partial E}{\partial \lambda}\right)_S \frac{d\lambda}{dt}\,.$$

As a result,

$$\frac{\overline{\partial \mathcal{H}(p,q;\lambda)}}{\partial \lambda} = \left(\frac{\partial E}{\partial \lambda}\right)_S\,.$$

4.3 Pressure

We have already mentioned that both the entropy and the energy are *additive quantities*, i.e. they are proportional to system size, and therefore do not depend on the *shape* for systems in thermal equilibrium. We may consider the entropy to be a function of energy, $S(E)$, or equivalently the energy to be a function of the entropy, $E(S)$, for the closed system, i.e. in the case when we have no external system or medium. Let us consider below the case where the system does not take part in a macroscopic motion. Then the energy E is equal to the *internal energy* which we will denote as U. Thus, in our case, $U = \overline{\mathcal{H}(p,q)}$.

Now, consider a process in which the wall of the container moves *adiabatic*, i.e. slowly enough for the system to stay in internal thermal equilibrium. A force $\mathbf{F}$ acts from the contents of the system on the container wall surface element, $\delta\mathbf{s}$. If the container wall moves a small distance, $\delta\mathbf{r}$, then work is done on the surroundings given by $\delta W = \mathbf{F}\cdot\delta\mathbf{r}$. This work must be equal to the change of the internal energy of the system. Consequently, the force acting on a surface element $d\mathbf{s}$ is

$$\mathbf{F} = -\frac{\overline{\partial \mathcal{H}(p,q;\mathbf{r})}}{\partial \mathbf{r}} = -\left(\frac{\partial U}{\partial \mathbf{r}}\right)_S = -\left(\frac{\partial U}{\partial V}\right)_S \frac{\partial V}{\partial \mathbf{r}} = -\left(\frac{\partial U}{\partial V}\right)_S d\mathbf{s}\,.$$

The magnitude of the force per area $d\mathbf{s}$ is called the pressure,

$$\mathbf{P} = -\left(\frac{\partial U}{\partial V}\right)_S\,. \tag{4.2}$$

Combining definitions of the temperature and pressure, one can write

$$dU = T\,dS - P\,dV\,. \tag{4.3}$$

This relation is often called *the first law of thermodynamics*. The pressures of the bodies in the equilibrium are the same (to preserve mechanical stability). [2]

It follows from Eq. (4.3) that

$$T = \left(\frac{\partial U}{\partial S}\right)_V, \quad P = -\left(\frac{\partial U}{\partial V}\right)_S\,. \tag{4.4}$$

---

[2] This property can be violated in *metastable* states.

The first law (4.3) can be rewritten as

$$dS = \frac{1}{T}dU + \frac{P}{T}dV\,, \tag{4.5}$$

and it follows that

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_U . \tag{4.6}$$

It the microscopic expression for the entropy is known, this equation allows one to calculate the *equation of state* connecting P, T and V.

4.4 Equations of state

As we have seen, the systems like liquid or gas in an equilibrium can be described by state variables like temperature T, pressure P, or total volume V. They are not independent since the variables in equilibrium related by an equation of state. Usually it is taken in the form,

$$P = P(V,T)\,. \tag{4.7}$$

Thus, only two variables can be assigned independently. The equation of state can be obtained either from experiment, or from microscopic theory.

An ideal gas of N particles obeys the equation of state

$$P = NT/V = \rho T \quad \text{(temperature is measured in energy units)} \tag{4.8}$$

Here $\rho \equiv N/V$ is the gas density.

For one mole of the gas N equals to the Avogadro's number $N_A = 6.023 \times 10^{23}$ mole^{-1}. If T is measured in Kelvin, then $T \to k_B T$, and the equation of state for 1 mole reads

$$P = RT/V\,, \quad R \equiv N_A k_B = 8.314\ \text{J}/(\text{K}\cdot\text{mole})\,.$$

Including *interaction* between particles in a simple approximate form leads to more realistic van der Waals equation,

$$P = \frac{NT}{V - bN} - a\left(\frac{N}{V}\right)^2 = \frac{N\rho}{1 - b\rho} - a\rho^2\,. \tag{4.9}$$

The item $b\rho$ in the denominator take care of particle repulsion at small distances, while the term $\propto \rho^2$ allows for particle attraction at relatively large distances.

The phase diagram of real gases can be obtained from the equation of state. In general, it has the form shown in Fig. 4.1. We will come back to phase transitions later where various phase diagrams will be discussed.

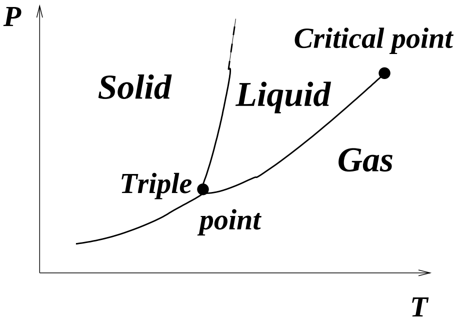

Figure 4.1: Phase diagram of a real liquid-gas system.

4.5 Work and quantity of heat

Consider a closed system that consists of two parts – the 'system' and the surrounding *medium* or '*heat bath*', Fig. 4.2

First consider the situation in which the system is thermally insulated from the medium. If the volume of the system changes by δV then the medium, which is has a pressure P (equal to the pressure of the medium), performs *work* on the system.

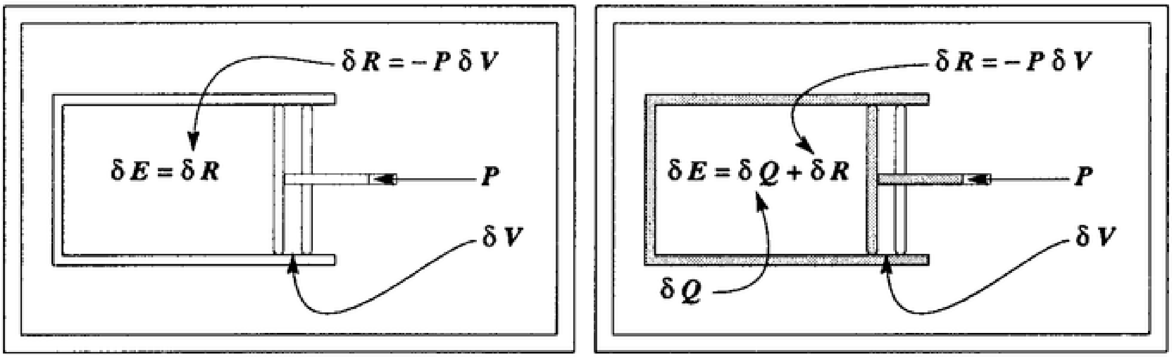

Figure 4.2: A closed system consisting of a heat bath surrounding the system of interest. (a) – The surroundings work on a thermally insulated system by changing its volume by δV. The work done by the system is $\delta W = -P\delta V$. The increase in energy of the system is $\delta U = \delta W = -P\delta V$. (b) – The walls of the system conduct *heat* and in the process of changing the volume of the system the heat δQ enters the system, therefore the energy change is $\delta U = \delta Q + \delta W$. Together the system and the surroundings form a closed system

We assume that $W > 0$ when external forces produce *work* on the body in question, while $W < 0$ if the body itself produces work. One obtains

$$dW/dt = -P\,dV/dt$$

when the volume changes in time. For a reversible process the pressure must be *the same* throughout the body.

If the body is *n*ot isolated thermally, then

$$\frac{dE}{dt} = \frac{dW}{dt} + \frac{dQ}{dt},$$

where Q is the *heat* gained or lost by the body per unit time. Here E is understood as a whole energy of the body including macroscopic motion. For the body at rest E equals to the *internal* energy U. In this case,

$$\frac{dQ}{dt} = \frac{dU}{dt} + P\frac{dV}{dt}.$$

Assuming that the body is in an equilibrium corresponding to its energy and the volume at the instant, one can put $U = U(S,V)$,

$$\frac{dU}{dt} = T\frac{dS}{dt} - P\frac{dV}{dt}.$$

Thus

$$\frac{dQ}{dt} = T\frac{dS}{dt}. \qquad (4.10)$$

Note that dQ and dW are not *complete differentials*, i.e. that they depend on the way in which the states of the system change in time.

One can imagine the situation when the pressure and temperature are constant throughout the body, while the state is non-equilibrium (say, chemical reaction in the mixture). Since this process is *irreversible*, one gets the inequality

$$\frac{dQ}{dt} < T\frac{dS}{dt}. \qquad (4.11)$$

instead of Eq. (4.10). If the systems goes from one equilibrium state to another equilibrium state through some non-equilibrium states, than

$$\delta Q \le T\delta S. \qquad (4.12)$$

This relation can be considered as a form of second law of thermodynamics.

4.6 Thermodynamics potentials

4.6.1 The heat function (enthalpy)

If the process occurs at constant volume, $dQ = dU$. If the *pressure* is kept constant,

$$dQ = d(U + PV) \equiv dH, \quad H(S,P) = U + PV. \qquad (4.13)$$

Since

$$dU = T\,dS - P\,dV, \quad dH = dU + P\,dV + V\,dP,$$

one obtains

$$dH = T\,dS + V\,dP\,. \tag{4.14}$$

Thus,

$$T = (\partial H/\partial S)_P\,, \quad V = (\partial H/\partial P)_S\,. \tag{4.15}$$

As a result, if the body is *thermally isolated and the pressure is kept constant*, then

$$H = \text{const}\,.$$

4.6.2 Helmholtz free energy and thermodynamic potential (Gibbs free energy)

Consider the work on a body in a reversible isothermal change of the state,

$$dW = dU - dQ = dU - T\,dS = d(U - TS) = dF\,, \quad F(V,T) \equiv U - TS\,. \tag{4.16}$$

F is called the (Helmholtz) free energy. Since

$$dU = T\,dS - P\,dV\,, \quad dF = dU - T\,dS - S\,dT$$

one obtains,

$$dF = -S\,dT - P\,dV\,. \tag{4.17}$$

Thus,

$$S = -(\partial F/\partial T)_V\,, \quad P = -(\partial F/\partial V)_T\,. \tag{4.18}$$

Since $U = F + TS$

$$U = F - T(\partial F/\partial T)_V = -T^2\left(\frac{\partial}{\partial T}\frac{F}{T}\right)_V\,.$$

Thus, if we know any of the quantities, $U(S,V)$, $F(T,V)$, or $H(S,P)$, we can determine all the remaining thermodynamic quantities. Thus U,F and H are called the thermodynamic potentials with respect to their variables.

We still miss the thermodynamic potential with respect to P,T. Introducing

$$G = U - TS + PV = F + TS = H - TS$$

and using the equality

$$P\,dV = d(PV) - V\,dP$$

we obtain

$$dG = -S\,dT + V\,dP \quad \to \quad S = -(\partial G/\partial T)_P,\; V = (\partial G/\partial P)_T\,. \tag{4.19}$$

If there are other parameters, λ_i the expression

$$\sum_i \Lambda_i\,d\lambda_i$$

should be added to all thermodynamic potentials. In this way we can obtain the expressions for the averages,

$$\Lambda_i = \overline{\frac{\partial \mathcal{H}(p,q;\lambda)}{\partial \lambda_i}} = \left(\frac{\partial F}{\partial \lambda_i}\right)_{T,V,\lambda_{j\neq i}} = \left(\frac{\partial H}{\partial \lambda_i}\right)_{S,P,\lambda_{j\neq i}} = \dots .$$

If the parameters λ_i change slightly, the changes in the thermodynamic potentials are equal if each is considered for the appropriate pair of constant quantities,

$$(\delta U)_{S,V} = (\delta F)_{T,V} = (\delta H)_{S,P} = (\delta G)_{T,P} . \tag{4.20}$$

This statement is called the theorem of small increments.

Since

$$\frac{dU}{dt} + P\frac{dV}{dt} < T\frac{dS}{dt}$$

at constant V and T we get $dF/dt < 0$, while at constant P and T $dG/dt < 0$. Thus at the equilibrium the thermodynamic potentials reach minima with respect to the variations of state with proper constant parameters.

4.6.3 Dependence on the number of particles

Since S and V are *additive* the internal energy can be written as

$$U = N f(S/N, V/N)$$

since this is the most general *homogeneous* function of the first order in N,S and V. In a similar way,

$$F = N f(V/N,T), \quad H = N f(S/N,P), \quad G = N f(P,T).$$

If one considers N as a formal independent variable, then all the thermodynamic potentials acquire additional terms $\mu\, dN$, where

$$\mu = \left(\frac{\partial U}{\partial N}\right)_{S,V} = \left(\frac{\partial H}{\partial N}\right)_{S,P} = \left(\frac{\partial F}{\partial N}\right)_{T,V} = \left(\frac{\partial G}{\partial N}\right)_{P,T} .$$

μ is called the chemical potential. We easily find that $G = N\mu(P,T)$. Thus

$$d\mu = -(S/N)\, dT + (V/N)\, dP .$$

If one considers the volume is constant, but the number of particles as a variable, the independent variables are, say, T,N. Then it is convenient to introduce the new thermodynamic potential,

$$\Omega = F - \mu N = F - G = -PV .$$

Then

$$d\Omega = -S\, dT - N\, d\mu .$$

Sometimes Ω is called the Landau free energy. The number of particles can be expressed through $\Omega(T,V,\mu)$ as

$$N = -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = V\left(\frac{\partial P}{\partial\mu}\right)_{T,V}.$$

One can show that if 2 bodies are brought together and they can exchange by particles, then

$$\mu_1 = \mu_2$$

We can summarize the thermodynamic potentials in the table 4.1.

Thermodynamic potential	Notation	Independent variables	Differential
Internal energy	U	$S,\ V,\ N$	$T\,dS - P\,dV + \mu\,dN$
Heat function (enthalpy)	H	$S,\ P,\ N$	$T\,dS + V\,dP + \mu\,dN$
Helmholtz free energy	F	$T,\ V,\ N$	$-S\,dT - P\,dV + \mu\,dN$
Gibbs free energy	G	$T,\ P,\ N$	$-S\,dT + V\,dP + \mu\,dN$
Landau free energy	Ω	$T,\ V,\ \mu$	$-S\,dT - P\,dV - N\,d\mu$

Table 4.1: Thermodynamic potentials (summary)

4.6.4 Derivatives of thermodynamic functions

The quantity of heat which must be gained in order to raise the temperature of the body by one unit is called the specific heat. We get for the cases of fixed volume and pressure, respectively:

$$C_V = T(\partial S/\partial T)_V\,, \quad C_P = T(\partial S/\partial T)_P\,.$$

Another important response functions are the compressibilities,

$$K_X = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_X \qquad (x = V, P,\ \text{or}\ S)\,,$$

and the thermal expansion coefficient

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

Usually, pairs of thermodynamic variables T,V and T,P are mostly used. If V and T are used as *independent variables*, the results can be expressed through the specific heat $C_V(V,T)$ and through pressure P. If we know the equation of state, $P(V,T)$, then we can calculate everything.

Similarly, if P and T are used as independent variables, the results can be expressed through the specific heat $C_P(P,T)$ and through the equation of state $V = V(P,T)$.

As an example of such derivation one can find the relation between C_P and C_V considering the internal energy U as a function T and V. Since

$$T\,dS = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[P + \left(\frac{\partial U}{\partial V}\right)_T\right] dV$$

we obtain assuming the pressure to be constant

$$C_P = C_V + \left[P + \left(\frac{\partial U}{\partial V}\right)_T\right]\left(\frac{\partial V}{\partial T}\right)_P . \tag{4.21}$$

Let us also demonstrate how one can obtain relations between second derivatives of thermodynamic quantities. Since

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

we have

$$\left(\frac{\partial C_V}{\partial V}\right)_V = T\frac{\partial^2 S}{\partial V \partial T} = -T\frac{\partial^3 F}{\partial V \partial T^2} = -T\frac{\partial^2}{\partial T^{2^2}}\left(\frac{\partial F}{\partial V}\right)_T .$$

Now, since

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

we get finally

$$\left(\frac{\partial C_V}{\partial V}\right)_T = T\left(\frac{\partial^2 P}{\partial T^2}\right)_V .$$

Consequently, this derivative can be calculated *from the equation of state*. Similarly,

$$\left(\frac{\partial C_P}{\partial P}\right)_T = T\left(\frac{\partial^2 V}{\partial T^2}\right)_P ,$$

that van be obtained using Gibbs free energy instead of the Helmholtz one (prove!).

Calculation of derivatives from the equation of state

The calculations are based on the so-called Maxwell relations. Consider a thermodynamic potential, $\mathcal{F}(X,Y)$, which depends on the independent variables X,Y. Then, $d\mathcal{F}$ can be written as

$$d\mathcal{F} = R_x\, dX + R_y\, dY .$$

Since

$$\frac{\partial^2 \mathcal{F}}{\partial X\, \partial Y} = \frac{\partial^2 \mathcal{F}}{\partial Y\, \partial X}$$

one obtains

$$\left(\frac{\partial R_x}{\partial Y}\right)_X = \left(\frac{\partial R_y}{\partial X}\right)_Y .$$

Example:

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V .$$

Other important relations (see derivation in Appendix A) arise from the properties of partial differentiations. Consider 3 quantities, X,Y,Z, related by the equation of state $\mathcal{K}(X,Y,Z) =$

const. We can consider X,Y as independent variables, while $Z=Z(X,Y)$. Another way is assume Y,Z to be independent variables, while $X=X(Y,Z)$. The relations between the derivatives of theses functions read

$$\left(\frac{\partial X}{\partial Y}\right)_Z \left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y = -1\,, \tag{4.22}$$

$$\left(\frac{\partial X}{\partial Y}\right)_Z = \left(\frac{\partial Y}{\partial X}\right)_Z^{-1}\,. \tag{4.23}$$

Any quantity $\mathcal{F}(X,Y)$, the differential of which can be expressed as

$$d\mathcal{F} = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_Y dX + \left(\frac{\partial \mathcal{F}}{\partial Y}\right)_X dX\,.$$

From this expression we get

$$\left(\frac{\partial \mathcal{F}}{\partial Z}\right)_Y = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_Y \left(\frac{\partial X}{\partial Z}\right)_Y\,, \tag{4.24}$$

and

$$\left(\frac{\partial \mathcal{F}}{\partial X}\right)_Z = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_Y + \left(\frac{\partial \mathcal{F}}{\partial Y}\right)_X \left(\frac{\partial Y}{\partial X}\right)_Z\,. \tag{4.25}$$

Equations (4.24), (4.25) together with Eqs. (4.22), (4.23) and the Maxwell relations are usually used for transformations and computation of derivatives from the equation of state.

4.6.5 Examples

Entropy:

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

Here we have used the realtion $dF = -S\,dT - P\,dV$. Similarly,

$$\left(\frac{\partial S}{\partial P}\right)_T = -\frac{\partial}{\partial P}\left(\frac{\partial G}{\partial T}\right)_P = -\frac{\partial}{\partial P}\left(\frac{\partial G}{\partial P}\right)_T = -\left(\frac{\partial V(P,T)}{\partial T}\right)_P\,.$$

Internal energy:

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial S(V,T)}{\partial V}\right)_T - P = T\left(\frac{\partial P(V,T)}{\partial T}\right)_V - P\,.$$

Here we have used the realtion $dU = T\,dS - P\,dV$. Similarly,

$$\left(\frac{\partial U}{\partial P}\right)_T = -T\left(\frac{\partial V(P,T)}{\partial T}\right)_P - P\left(\frac{\partial V(P,T)}{\partial P}\right)_T\,,$$

$$\left(\frac{\partial U}{\partial T}\right)_P = C_P - P\left(\frac{\partial V}{\partial T}\right)_P .$$

One can also obtain the following relations (prove!)

$$\begin{aligned} C_P - C_V &= (TV/K_T)\alpha^2 ; \quad (4.26)\\ K_P - K_V &= (TV/C_P)\alpha^2 ; \quad (4.27)\\ C_P/C_V &= K_T/K_S . \quad (4.28)\end{aligned}$$

4.7 Principles of thermodynamic

4.7.1 Maximum work and Carnot cycle

Consider a thermally isolated system consisting of several bodies not in thermal equilibrium, see Fig. 4.3. The system may do during establishment of equilibrium. This work, as well as the final equilibrium state, depend on the concrete way which the system follows.

Let the total original energy be E_0 while the energy in the final equilibrium state is $E(S)$. Since the system is thermally isolated,

$$|W| = E_0 - E(S), \quad \partial|W|/\partial S = (\partial E/\partial S)_V = -T .$$

Here T is the temperature in the final state. Since the derivative is negative, the work decreases with increasing S. Since the entropy cannot decrease the greatest possible work is done when the entropy retains constant throughout the process. That means the reversible way.

Now let us consider two bodies with different temperatures, $T_2 > T_1$. If one brings the bodies into thermal contact, no work will be done since the process is irreversible. The entropy of two bodies will increase be $(\delta E)(1/T_1 - 1/T_2)$ where δE is the energy transfer. To get a maximum work one has to accomplish a reversible cyclic process in a way to keep the bodies among which the direct energy transfer takes place at the same temperature. For that purpose one needs a further body (working medium). The working medium at temperature T_2 is brought in contact with the "heater" and receives some energy isothermal. It is then adiabatic cooled at the temperature T_1 and then adiabatic returned to the original state. The cyclic process is called the *Carnot cycle*.

Carnot discussed a process that has four stages:

1. The heat engine, at temperature T_2, is brought into contact with the reservoir at T_2, and receives the amount $-\delta E_2$ *isothermally*.

2. The heat-engine is removed from the T_2-reservoir, thermally isolated and undergoes an *adiabatic expansion* until its temperature is decreased to T_1.

3. The heat-engine is put in thermal contact with the T_1 reservoir, and gives off isothermally the energy δE_1 at temperature T_1.

4. The heat-engine is removed from the T_1-reservoir, thermally isolated and undergoes an *adiabatic compression* until its temperature is increased to T_2. The cycle is then, and the four stages may be repeated.

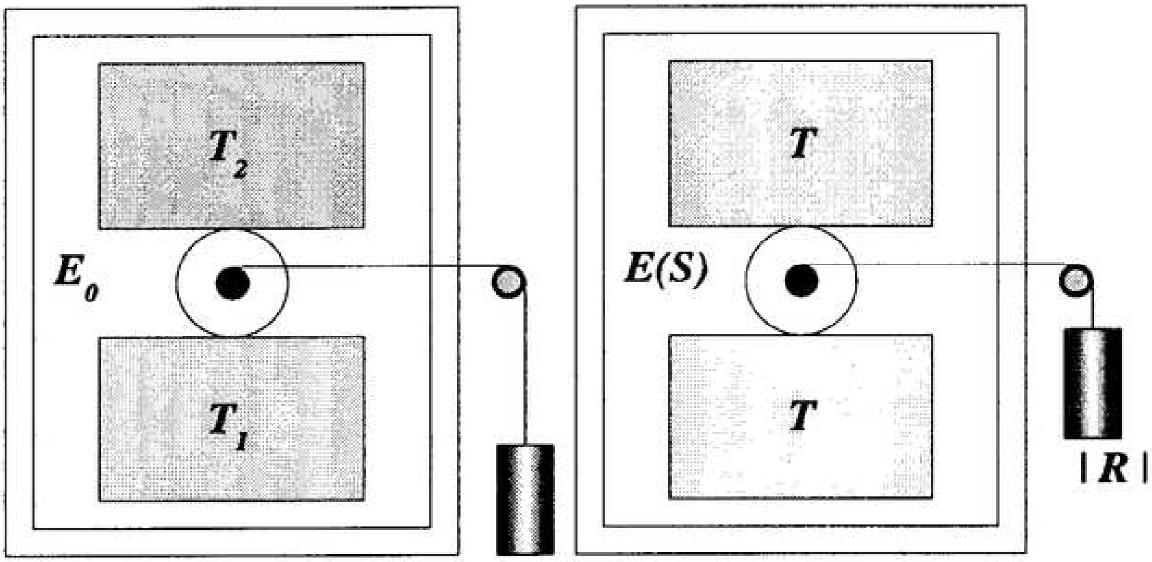

Figure 4.3: A thermally isolated system contains two parts at different temperatures $T_2 > T_1$. The two parts are connected via a heat engine that may perform the work $|W|$ on the surroundings outside the system. as the system approaches equilibrium. (a) – Initially the energy of the system is E_0. (b) – Finally the two parts have the same temperature T, and an amount of work was done on the surroundings. The energy of the equilibrium system is a unique function of entropy, $E(S)$. Both E and S depends on how the equilibrium state was reached.

The heat-engine performs work on the outside world during expansion, and receives work during contraction.

To calculate the work one can ignore the working medium for which the initial and final states are the same. Since

$$-\delta U_2 = -T_2\,\delta S_2\,, \quad \delta U_1 = T_1\,\delta S_1\,, \quad \delta S_2 + \delta S_1 = 0 \text{ (reversible process)}$$

we obtain

$$|\delta W|_{\max} = -\delta U_1 - \delta U_2 = -T_1\,\delta S_1 - T_2\,\delta S_2 = -(T_2 - T_1)\,\delta S_2 = \frac{T_2 - T_1}{T_2}\,|\delta U_2|,.$$

Consequently, the *maximum efficiency* $\eta \equiv |W|/(\delta U_2)$ is

$$\eta_{\max} = (T_2 - T_1)/T_2 < 1\,.$$

Lord Kelvin realized that the Carnot cycle may be used to define an *absolute scale of temperature* – the scale used here. Since the efficiency $\eta_{\max} \leq 1$ it follows that there is a lower limit of (equilibrium) of zero. The Kelvin scale therefore has a natural fixed point, absolute zero. Since $\eta_{\max}$ depends only on the ratio T_1/T_2, there is an arbitrary scale factor that is fixed by setting $T = 273.16$ K for the triple point of pure water. With this scale one degree on the Celsius scale equals one degree on the Kelvin scale.

Work in an external medium

Consider a body embedded into a large external medium having $T_0 \neq T$, $P_0 \neq P$ where T, P are the temperature and the pressure of the body. Since the medium is large T_0 and P_0 are assumed to be fixed. We are interested in the maximal work $|W|_{\max}$ done by the thermally isolated body. This work is equal to the *minimal* work $W_{\min}$ done by an external source on the body. The total change of the internal[3] energy, ΔU consists of the three parts

(i) mechanical work W done by the external source;

(ii) the work $P_0\,\Delta V_0$ done by the medium;

(iii) the heat $-T_0\,\Delta S_0$ gained from the medium.

Thus

$$\Delta U = W + P_0\,\Delta V_0 - T_0\,\Delta S_0\,.$$

Since

$$\Delta V_0 + \Delta V = 0, \quad \Delta S_0 + \Delta S \geq 0$$

$$W \geq \Delta U - T_0\,\Delta S + P_0\,\Delta V\,.$$

For a reversible process,

$$W_{\min} = -|W|_{\max} = \Delta(U - T_0 S + P_0 V)\,. \tag{4.29}$$

We come to the following conclusions.

(i) If the volume and the temperature of the body remains constant and equal to T_0, then $W_{\min} = \Delta F$.

(iii) If the pressure and the temperature of the body remains constant and equal to P_0 and T_0, respectively, then $W_{\min} = \Delta G$.

(iii) If no mechanical work is done and the body is left itself, the irreversible processes lead to the inequality

$$\Delta(U - T_0 S + P_0 V) \leq 0\,.$$

One can relate the minimal work $W_{\min}$ with the deviation of the total entropy S_t from its equilibrium value $S_t(U_t)$. Indeed,

$$\Delta S_t = -\frac{dS_t(U_t)}{dU_t} W_{\min} = -\frac{\Delta U - T_0\,\Delta S + P_0\,\Delta V}{T_0} = -\frac{(T - T_0)\Delta S - (P - P_0)\Delta V}{T_0}\,.$$

To reach maximum entropy one should require $T = T_0$, $P = P_0$.

---

[3] We assume that there is no macroscopic motions as a whole.

4.7.2 Thermodynamic inequalities. Thermodynamic stability

The vanishing of the 1st derivatives of thermodynamic quantities, though provide conditions for the equilibrium, like $T_1 = T_2$, $P_1 = P_2$, is only a necessary condition for an extremum and do not ensure that the entropy reaches the maximum. The sufficient conditions require analysis the second derivatives.

As we have stated, the quantity

$$\Delta U - T_0 \, \Delta S + P_0 \, \Delta V$$

has a minimum. Thus

$$\delta U - T_0 \, \delta S + P_0 \, \delta V > 0$$

for any small deviation from the equilibrium. Thus, the minimum work which must be done to bring the selected part from equilibrium to some other state is positive.

Now we can expand the internal energy $U(S,V)$ in powers of the variables,

$$\delta U = \frac{\partial U}{\partial S}\delta S + \frac{\partial U}{\partial V}\delta V + \frac{1}{2}\left[\frac{\partial^2 U}{\partial S^2}(\delta S)^2 + \frac{\partial^2 U}{\partial S \partial V}\delta S \delta V + \frac{\partial^2 U}{\partial V^2}(\delta V)^2\right].$$

The 1st order terms provide $T\,\delta S - P\,\delta V$, and we have

$$\frac{\partial^2 U}{\partial S^2}(\delta S)^2 + \frac{\partial^2 U}{\partial S \partial V}\delta S \delta V + \frac{\partial^2 U}{\partial V^2}(\delta V)^2 > 0\,.$$

Since δS and δV are *arbitrary*, one has to assume

$$\frac{\partial^2 U}{\partial S^2} > 0\,, \tag{4.30}$$

$$\frac{\partial^2 U}{\partial V^2} > 0\,, \tag{4.31}$$

$$\frac{\partial^2 U}{\partial S^2}\frac{\partial^2 U}{\partial V^2} - \left(\frac{\partial^2 U}{\partial S \partial V}\right)^2 > 0\,. \tag{4.32}$$

Since

$$\frac{\partial^2 U}{\partial S^2} = \left(\frac{\partial T}{\partial S}\right)_V = \frac{T}{C_V}$$

we obtain from Eq. (4.30)

$$C_V > 0\,.$$

In a similar way,

$$\frac{\partial^2 U}{\partial V^2} - = -\left(\frac{\partial P}{\partial V}\right)_S = \frac{1}{VK_S} > 0\,.$$

From Eq. (4.32) one obtains

$$\left(\frac{\partial T}{\partial V}\right)_S^2 < \frac{T}{VK_S C_V}\,.$$

After some algebra (see the book by Landau and Lifshitz, §21) one can get from Eq. (4.32) the inequality

$$\left(\frac{\partial P}{\partial V}\right)_T < 0 .$$

Here we have calculated the conditions for thermodynamic stability.

4.7.3 Nernst's theorem

As we know, for normal systems $S \to 0$ at $T \to 0$. This statement is called the Nernst's theorem. If it is true,

$$C_P = C_V = 0 \quad \text{at} \quad T = 0 .$$

One can obtain similar equalities for $T = 0$,

$$\begin{aligned} (\partial V/\partial T)_P &= -(\partial S/\partial P)_T = 0 \quad \text{(Maxwell relation for } G\text{)} \\ (\partial P/\partial T)_V &= 0 \\ (C_P - C_V)/C_P &= 0 . \end{aligned}$$

The following useful formulas follow from the Nernst's theorem:

$$\begin{aligned} S &= \int_0^T (C_V/T)\, dT , \\ H &= H_0 + \int_0^T C_P\, dT , \\ G &= H_0 + \int_0^T C_P\, dT - T \int_0^T (C_V/T)\, dT . \end{aligned}$$

4.7.4 On phase transitions

The equilibrium between the phases requires

$$T_1 = T_2 , \quad P_1 = P_2 , \quad \mu_1 = \mu_2 .$$

Along the the co-existence curve between gas and liquid,

$$\mu_g(T,P) = \mu_l(T,P) , \quad d\mu_g(T,P) = d\mu_l(T,P) .$$

Introducing entropy per particle $s \equiv S/N$ and volume $v \equiv V/N$ we get

$$-s_g\, dT + v_g\, dP = -s_l\, dT + v_l\, dP .$$

Consequently,

$$\left(\frac{dP}{dT}\right)_{\text{coexist}} = \frac{s_g - s_l}{v_g - v_l} = \frac{\lambda}{T(v_g - v_l)} ,$$

where $\lambda \equiv T(s_g - s_l)$ is the latent heat. This is the *Clausius-Clapeyron equation*. From the equations of state for gas and liquid phases one can find the dependence $P(T)$ along the coexistence line. The latent heat λ depends on the temperature and vanishes at some temperature T_c. At $T = T_c$ the two phases become identical. So the entropy changes continuously while its first derivatives change abruptly. We shall come back to phase transitions later in much greater detail.

Chapter 5

The Gibbs Distribution

5.1 Spin $\frac{1}{2}$ in a magnetic field

Let us start with a simple example. Consider N noninteracting *localized* particles with spin $S=\frac{1}{2}$ embedded in an external magnetic field $\mathbf{H}$. The Hamiltonian is

$$\varepsilon = -\mathbf{m}\cdot\mathbf{H} = -m_z H\,, \quad \mathbf{m} = 2\mu\mathbf{S}\,,$$

$\mu = e/2mc$ is the Bohr magneton. Since $S_z = \pm\frac{1}{2}$,

$$\varepsilon_{\uparrow(\downarrow)} = \mp\mu H\,.$$

Denoting $n_{\uparrow(\downarrow)}$ as numbers of states with spin up(down), respectively, we get for the energy

$$E = -\mu(n_\uparrow - n_\downarrow)\,H\,, \tag{5.1}$$

with the normalization condition

$$N = n_\uparrow + n_\downarrow\,. \tag{5.2}$$

The total magnetic moment is

$$M = \mu(n_\uparrow - n_\downarrow)\,.$$

Under the assumptions of the *microcanonical ensemble* (i. e. with fixed total moment M), the thermodynamic weight for the system is

$$W = \frac{N!}{n_\uparrow!n_\downarrow!}\,.$$

The numbers $n_{\uparrow(\downarrow)}$ can be expressed in terms of E and M. Using the Stirling's formula, [1]

$$\log n! \approx n\log n - n + 1 \quad \text{at} \quad n \gg 1 \tag{5.3}$$

---

[1] When $N \gg 1$ one can replace $\sum_n = 0^N \ln n$ as $\int_0^N \ln x\, dx = N\ln(N/e)$.

we obtain

$$S = N\log N - n_\uparrow \log n_\uparrow - n_\downarrow \log n_\downarrow\,.$$

Now we can express the quantities $n_{\uparrow(\downarrow)}$ in terms of the energy E given by Eq. (5.1) as

$$n_{\uparrow(\downarrow)} = \frac{\mu HN \mp E}{2\mu H}\,.$$

consequently,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_N = \frac{1}{2\mu H}\log\frac{\mu HN + H}{\mu HN - H} = \frac{1}{2\mu H}\log\frac{n_\uparrow}{n_\downarrow}\,. \tag{5.4}$$

Thus,

$$M = N\mu\tanh\frac{\mu H}{T}\,, \qquad \chi \equiv \left(\frac{\partial M}{\partial H}\right)_T = \frac{N\mu^2}{T}\frac{1}{\cosh^2(\mu H/T)}\,.$$

χ is called the magnetic susceptibility.

Now let us define the same relations from the *canonical ensemble*, i. e. assuming that the temperature is fixed by a thermal contact with a “thermal bath”. Since the energy cost for one spin flip

$$n_\uparrow \to n_\uparrow - 1\,, \quad n_\downarrow \to n_\downarrow + 1$$

is $\Delta U = 2\mu H$ and the entropy change is

$$\Delta S = \log\frac{N!}{(n_\uparrow - 1)!(n_\downarrow + 1)!} - \log\frac{N!}{n_\uparrow! n_\downarrow!} = \log\frac{n_\uparrow}{n_\downarrow + 1}$$

we obtain

$$\frac{\Delta S}{\Delta U} \approx \frac{1}{2\mu H}\log\frac{n_\uparrow}{n_\downarrow}\,.$$

That must be equal to the inverse temperature, and we recover previous results.

The previous results can be summarized as a distribution function

$$n_\sigma = \frac{N}{Z_1}e^{-\beta\varepsilon_\sigma}\,, \quad Z_1 = \sum_{\sigma=\uparrow,\downarrow} e^{-\beta\varepsilon_\sigma} = 2\cosh\frac{\mu H}{T}\,. \tag{5.5}$$

Here $\beta = 1/T$. We have

$$\bar{\varepsilon} = \frac{1}{Z_1}\sum_{\sigma=\uparrow,\downarrow}\varepsilon_\sigma e^{-\beta\varepsilon_\sigma} = -\mu H\tanh\frac{\mu H}{T}\,, \quad U = N\bar{\varepsilon}.$$

Now there exist *fluctuations* around the average energy which are absent in the microcanonical ensemble.

Exercise: Calculate the fluctuation in energy, $\sqrt{\overline{\varepsilon^2} - E^2}$ and compare it with the internal energy $E = N\bar{\varepsilon}$. Show that the difference is small for a macroscopic body.

5.2 The Gibbs distribution

Eq. (5.5) is the particular case of the Gibbs distribution. If there is a body, embedded into a medium, the microcanonical distribution can be written as

$$dw \propto \delta(E+E'-E^{(0)})\, d\Gamma\, d\Gamma'$$

where superscript "prime" relate the quantities to the medium. If one fixes the body in a given state n, then the probability to find the body in this state is

$$f_n \propto \int \delta(E_n+E'-E^{(0)})\, d\Gamma' = \int \delta(E_n+E'-E^{(0)})\, \frac{d\Gamma'}{dE'}\, dE' = \int \frac{e^{S'(E')}}{\Delta E'} \delta(E_n+E'-E^{(0)})\, dE' \,.$$

As a result,

$$f_n \propto \left(\frac{e^{S'(E')}}{\Delta E'}\right)_{E'=E^{(0)}-E_n} .$$

Now let us consider the body as very small comparing to the medium. Then

$$E_n \ll E^{(0)} \to \quad S'(E^{(0)}-E_n) \approx S'(E^{(0)}) + E_n \frac{dS'(E^{(0)})}{dE^{(0)}} = S'(E^{(0)}) - E_n/T \,.$$

Finally,

$$f_n = \frac{1}{Z} e^{-\beta E_n}, \qquad Z = \sum_s e^{-\beta E_s}, \qquad \beta \equiv 1/T \,.$$

In a classical system,

$$f(p,q) = \frac{1}{Z} e^{-\beta E(p,q)}, \qquad Z = \int e^{-\beta E(p,q)}\, d\Gamma \,.$$

Note that in the derivation we use the Hamiltonian of the whole system.

The Gibbs distribution is relevant for a body in an equilibrium with a large medium - "thermal bath".

If the systems consists of non-interacting particles which have eigenenergies ε_s, sth state being degenerate by g_s times, then we arrive at the single-particle distribution function,

$$f_n = \frac{g_n}{Z_1} e^{-\beta \varepsilon_n}, \qquad Z_1 = \sum_s g_s e^{-\beta \varepsilon_s} \,.$$

Here Z_1 is the *single-particle partition function.*

The Maxwell-Boltzmann distribution

Derivation from the Gibbs distribution

Consider a system of non-interacting particles. Dividing the total energy into potential and kinetic parts,

$$E(p,q) = V(q) + K(p)$$

we can write

$$\begin{aligned} f(p,q) &= f_q(q)\cdot f_p(p) \\ f_q &= e^{-\beta V(q)}/\left(\int dq\, e^{-\beta V(q)}\right); \\ f_p &= e^{-\beta K(p)}/\left(\int dp\, e^{-\beta K(p)}\right). \end{aligned}$$

For non-interacting particles the single-particle distribution is

$$f_p(p) = (2\pi m T)^{-3/2}\, e^{-(p_x^2+p_y^2+p_z^2)/2mT}\,.$$

Exercise: Show that the average kinetic energy per atom is $\frac{3}{2}T$.

Statistical derivation

Consider particles with energy levels ε_i and degeneracy factor g_i, $N=\sum_i n_i$. Here n_i is the number of particle occupying ith level. If the particles are *identical* and *distinguishable* the total number of microstates for a given state specified by the set $\{n_i\}$ is

$$W = N!\prod_i \frac{g_i^{n_i}}{n_i!}\,, \quad E=\sum_i n_i\varepsilon_i\,.$$

The easiest way is to use canonical ensemble which specifies the temperature. Let us consider the transition from jth to ith state,

$$n_i \to n_i+1\,, \quad n_j \to n_j-1\,.$$

Then

$$\Delta U = T\,\Delta S = \varepsilon_i - \varepsilon_j\,, \tag{5.6}$$

while

$$\Delta S = \log \frac{g_i^{n_i+1}}{(n_i+1)!}\frac{g_j^{n_j-1}}{(n_j-1)!} - \log \frac{g_i^{n_i}}{n_i!}\frac{g_j^{n_j}}{n_j!} = \log \frac{g_i n_j}{g_j(n_i+1)}\,. \tag{5.7}$$

Comparing Eqs. (5.6) and (5.7), we get

$$\frac{g_i n_j}{g_j(n_i+1)} = e^{\beta(\varepsilon_i-\varepsilon_j)}\,.$$

Thus we arrive at the same result with $Z=Z_1=\sum_i g_i\, e^{-\beta\varepsilon_i}$ where ε_i are *single-particle* energy levels.

Expressions for thermodynamic quantities

Average single-particle energy:

$$\bar{\varepsilon} = \frac{1}{Z_1}\sum_i \varepsilon_i g_i e^{-\beta\varepsilon_i} = -\frac{\partial \log Z_1(\beta)}{\partial\beta}, \quad U = N\bar{\varepsilon}.$$

Entropy: Since the entropy is given by the relation

$$S = -\langle \log f_n \rangle = -\sum_n f_n \log f_n$$

we obtain in general

$$S = \log Z + U/T\,.$$

The same result can be obtained from single-particle considerations. For N *distinguishable* particles we have

$$\begin{aligned} S &= \log W \approx N\log N - N + \sum_i (n_i \log g_i - n_i \log n_i) \\ &\approx N\log N - \sum_i n_i \log\frac{n_i}{g_i} = N\log N - \sum_i n_i \left(\log\frac{N}{Z_1} - \beta\varepsilon_i\right) \\ &= N\log Z_1 + U/T = \log Z_1^N + U/T = \log Z + U/T\,. \end{aligned}$$

Here we employ the fact that $Z = Z_1^N$. If the particles are *non-distinguishable,* them one has to divide the thermodynamic weight by $N!$ to obtain

$$Z = \frac{Z_1^N}{N!}, \quad S = -\sum_s g_s (f_s \log f_s - f_s)\,.$$

Helmholtz free energy:

$$F = U - TS = -T\log Z = -T\,\log\frac{Z_1^N}{N!}$$

This formula is fundamental in thermodynamic applications of the Gibbs distribution. Using this formula one can express the Gibbs distribution in the form

$$f_n = e^{-\beta(F-E_n)}\,.$$

5.2.1 The Gibbs distribution for a variable number of particles

.

If the body can exchange by particles with the medium, one has to specify the particle numbers, N and N', respectively, so that $N + N' = N^{(0)}$. Both the entropies and energies become functions on N, N'. As a result, we get

$$f_{nN} \propto \exp\left[S'(E^{(0)} - E_{nN}, N^{(0)} - N)\right]\,.$$

Next, we expand the entropy up to the 1st order in E_{nN} and N. Since

$$\left(\frac{\partial S}{\partial U}\right)_{V,N} = \beta\,, \quad \left(\frac{\partial S}{\partial N}\right)_{U,V} = -\beta\mu\,,$$

we get

$$S'(E^{(0)} - E_{nN}, N^{(0)} - N) = S'(E^{(0)}, N^{(0)}) - \beta E_{nN} + \beta\mu N$$

(due to the assumption of the equilibrium, the temperatures and chemical potentials of the body and the medium are the same). Thus,

$$f_{nN} = A e^{\beta(\mu N - E_{nN})}\,.$$

The normalization constant A can be expressed through the thermodynamic quantities. Since

$$S = -\langle \log f_{nN} \rangle = -\log A - \beta\mu\bar{N} + \beta\bar{E} \quad \rightarrow \quad T\log A = U - TS - \mu\bar{N} = \Omega\,,$$

we obtain

$$f_{nN} = e^{\beta(\Omega + \mu N - E_{nN})}\,.$$

Since

$$\sum_N \sum_s f_{sN} = e^{\beta\Omega} \sum_N \left(e^{\beta\mu N} \sum_s e^{\beta E_{sN}} \right) = 1\,,$$

one obtains

$$\Omega = -T \log \sum_N \left(e^{\beta\mu N} \sum_s e^{\beta E_{sN}} \right)\,.$$

Remember that for classical system of N non-distinguishable particles one has to use the classical Hamiltonian $\mathcal{H}(\{p, q; N\})$ instead E_{nN}, the phase volume element being

$$\frac{1}{(2\pi\hbar)^{dN} N!} \prod_{i=1}^{N} dp_i\, dq_i\,,$$

while for distinguishable particles there is no factor $1/N!$. Here d is the dimensionality of the real space.

Chapter 6

Ideal gas

6.1 Classical gas

Having in mind quantization in a box,

$$\psi_{\mathbf{n}}(\mathbf{r}) = \sin(k_x x)\cdot\sin(k_y y)\cdot\sin(k_z z)\,, \qquad k_i = \pi n_i/L \quad (n_i = 1,2,\ldots)$$

we obtain

$$\varepsilon_{\mathbf{n}} = \frac{\hbar^2\pi^2}{2mL^2}(n_x^2+n_y^2+n_z^2)\,.$$

Thus,

$$Z_1 = \sum_{n_x,n_y,n_z} e^{-(\beta\hbar^2\pi^2/2mL^2)(n_x^2+n_y^2+n_z^2)} \approx \left(\frac{1}{2}\int_{-\infty}^{\infty} dn_x\, e^{-(\beta\hbar^2\pi^2/2mL^2)n_x^2}\right)^3 = V/\Lambda^3\,. \tag{6.1}$$

Here

$$\Lambda = \left(2\pi\beta\hbar^2/m\right)^{1/2} = \hbar\sqrt{2\pi/mT} \propto T^{-1/2}$$

is the thermal de Broglie wave length. Another way to obtain this result is to use the continuous limit one can write

$$Z_1 = V\int \frac{d^3p}{(2\pi\hbar)^3} e^{-\beta\mathbf{p}^2/2m}\,.$$

It is very convenient to introduce the *density of states* ,

$$D(\varepsilon) = g\int \frac{d^3p}{(2\pi\hbar)^3}\delta\left(\varepsilon-\frac{\mathbf{p}^2}{2m}\right) = g\frac{(2m)^{3/2}}{4\pi^2\hbar^3}\sqrt{\varepsilon} \equiv D_0\sqrt{\varepsilon}\,, \quad D_0 \equiv g\frac{(2m)^{3/2}}{4\pi^2\hbar^3}\,. \tag{6.2}$$

Here g is the degeneracy factor of the state with given momentum $\mathbf{p}$ Now, the integral is easily done, to get

$$Z_1 = VD_0\sqrt{T}\int_0^\infty \sqrt{x}\,e^{-x}\,dx = V/\Lambda^3\,.$$

Introduction of the density of states helps in calculation for different systems.

Exercise: Calculate the partition function for two-dimensional gas.

Thermodynamic quantities

Now we proceed with thermodynamics:

$$\begin{aligned}
F &= -T\log\frac{Z_1^N}{N!} = -NT\log(V/\Lambda^3) + T\log N!\,; &(6.3)\\
P &= \left(\frac{\partial F}{\partial V}\right)_{T,N} = \frac{NT}{V} = \rho T \quad (\rho \text{ is the gas density})\,; &(6.4)\\
\mu &= \left(\frac{\partial F}{\partial N}\right)_{T,V} \approx -T\log\frac{Z_1}{N} = T\log\frac{P\Lambda^3}{T}\,; &(6.5)\\
U &= -\frac{\partial}{\partial\beta}\log Z = \frac{3}{2}NT\,; &(6.6)\\
C_V &= \frac{3}{2}N\,; &(6.7)\\
S &= -\left(\frac{\partial F}{\partial T}\right)_{V,N} = \log\frac{Z_1^N}{N1} + \frac{U}{T} = N\left[\frac{5}{2} - \log\rho\Lambda^3\right]. &(6.8)
\end{aligned}$$

Something is wrong: the entropy and the specific heat does not vanish at zero temperature. Indeed, the classical statistics is applicable only at

$$\rho\Lambda^3 \ll 1\,. \tag{6.9}$$

This inequality is violated at low temperatures. What was wrong in out derivation? We implicitly assumed that one state can either be empty, or occupied only by one particle. Indeed, according to Eq. (6.9), the number of states V/Λ^3 is much greater than the number of particles, N. Only because of this assumption the equality

$$Z = \frac{Z_1^N}{N!}$$

holds.

If the particles have internal degrees of freedom, k, with the energies ε_k, then the free energy acquires an additional item,

$$\delta F = -TN\log\sum_k e^{-\beta\varepsilon_k}\,.$$

Ideal gas with a constant specific heat

In many important cases the specific heat can be regarded as temperature-independent. Using the Stirling's formula for large number of particles , $N \gg 1$, we can write the Helmholtz free energy as

$$F = -NT\log\left[\frac{eV}{N\Lambda^3}\sum_k e^{-\beta\varepsilon_k}\right] \equiv -NT\log\frac{eV}{N} + Nf(T)\,, \quad C_V = -NTf''(T)\,,$$

the assumption $C_V = \text{const}(T)$ leads to the expression

$$f(T) = -c_V T\,\log T - \zeta_c T + \varepsilon_0\,,$$

where $c_V \equiv C_V/N$, while ζ_c and ε_0 are integration constants. ζ_c is called the *chemical constant*. We get,

$$\begin{aligned}
U &= N\varepsilon_0 + Nc_V T\,;\\
G &= N\varepsilon_0 + NT\,\log P - Nc_V T\,\log T - N\zeta_c T \quad (\text{since } PV = NT)\,;\\
H &= U + PV = N\varepsilon_0 + Nc_P T\,;\\
S &= N\,\log(eV/N) + Nc_V\,\log T + (\zeta_c + c_V)T\\
&= -N\,\log P + Nc_P\,\log T + (\zeta_c + c_P)N\,.
\end{aligned}$$

As a consequence, we can derive the *Poisson adiabatic*, i. e. the relation between V, P, and T for adiabatic expansion of compression of ideal gas with constant specific heat. Since

$$-N\,\log P + Nc_P\,\log T = \text{const}$$

one obtains

$$T^{c_P}/P = \text{const} \quad \to \quad T^\gamma P^{1-\gamma} = \text{const}\,, \quad \gamma = c_P/c_V\,.$$

Here we employed the relation $c_P - c_V = 1$ (prove!) which is valid for the ideal gas. Since $PV = NT$

$$TV^{\gamma-1} = \text{const}\,, \quad PV^\gamma = \text{const}\,.$$

Polyatomic gases

Vibrational degrees of freedom

If any internal motion of the molecules is present, then the potential energy of the molecule potential energy can be expanded in powers of the mechanical displacements q from the equilibrium positions,

$$V = \varepsilon_0 + \sum_{i,k=1}^{r_v} a_{ik} q_i q_k\,.$$

Here r_v is the number of *vibrational* degrees of freedom. If the molecule contains n atoms, then the total number of degrees of freedom is $3n$. If the atoms in the molecule are *collinear*, there are 3 translational degrees of freedom (for a molecule as a whole) and 2 rotational degrees of freedom, while $3n-5$ modes are left for vibrational motion. If the molecule is not a collinear one, then $r_v = 3n-6$. The kinetic energy depends on all $3n$ momenta. Consequently, the total energy and be written as

$$\varepsilon = \varepsilon_0 + f_{II}(p,q)\,,$$

where $f_{II}(p,q)$ is a quadratic function of $3n$ momenta and $3n-6$ (or $3n-5$) vibrational coordinates. Thus we have the total number of variables $l = 6n-6$, or $l = 6n-5$. For a monoatomic gas the total number of variables is 3 because coordinates do not appear in the expression for the energy. As a result, we obtain,

$$F = -NT \log \left[\frac{e \cdot e^{-\varepsilon_0/T}}{N} \int d\Gamma \, e^{-f_{II}(p,q)} \right] .$$

Now we can transform the variables as

$$p = p'\sqrt{T},\ q = q'\sqrt{T} \quad \to \quad f_{II}(p,q) = T\, f_{II}(p',q'),\ d\Gamma = T^{l/2} d\Gamma' .$$

In this way, we get

$$F = -NT \log \left(AV^{-\beta\varepsilon_0} T^{l/2}/N \right) .$$

Expanding the logarithm, we get

$$c_V = l/2 , \quad c_P = c_V + 1 = (l+2)/2 .$$

Thus *a classical gas has a constant specific heat*. We can formulate a rule:

Each degree of freedom contributes $T/2$ (or $kT/2$ in ordinary units) to the average energy.

Monoatomic ideal gas

Let exclude the kinetic energy of translational motion and label atomic *intrinsic levels* as ε'_k. Denoting

$$\mathcal{Z} = \sum_k e^{-\beta\varepsilon'_k}$$

we obtain

$$F = -NT \log \left[\frac{eV}{N\Lambda^3} \mathcal{Z} \right] = -NT \log \left[\frac{eV}{N} \left(\frac{mT}{2\pi\hbar^2} \right)^{3/2} \mathcal{Z} \right] .$$

In the simplest case, when the atoms are in the ground state with zero angular moment and spin, $L = S = 0$, then $\mathcal{Z} = e^{-\beta\varepsilon_0}$ where ε_0 is the energy of non-degenerate ground state. Starting the energy scale from ε_0 we obtain:

$$\mathcal{Z} = 1, \quad c_V = 3/2, \quad \zeta_c = (3/2)\log(m/2\pi\hbar^2) .$$

These results were obtained by Sackur and Tetrode in 1912. If only one of the angular momenta is non-zero in th ground state of the atom, then the situation is rather simple. Usually, in a ground state $L = 0$. while the total spin can be finite (e. g. vapor of alkali metals). Then $\mathcal{Z} = (2S+1)$ instead of 1, and $\zeta_S = \log(2S+1)$. If both momenta are non-zero than there is a fine structure in the levels which can be compared with the temperature. If we label the energy by the total angular momentum J, then

$$\mathcal{Z} = \sum_J (2J+1) e^{-\beta\varepsilon_J}$$

since each level with a given J is $(2J+1)$ degenerate with respect to directionds of J. This expression can be simplified at high temperatures, $\varepsilon_J \ll T$, when $e^{-\beta\varepsilon_J} \approx 1$. Then $\sum_J (2J+1) = (2L+1)(2S+1)$ and

$$\zeta_{SL} = \log[(2S+1)(2L+1)] .$$

Diatomic gases

In this cases additional rotational and vibrational degrees of freedom must be taken into account. Under some approximation, the energy can be written as sum to electronic, rotational, and vibrational contributions. In the simplest case of a singlet electronic state

$$\varepsilon_{vK} = \varepsilon_0 + \hbar\omega_0 \left(v + \frac{1}{2}\right) + \frac{\hbar^2}{2I} K(K+1) .$$

Here $\hbar\omega_0$ is the vibrational quantum, v is the vibrational quantum number, K is the rotational quantum number, $I = r_0^2 m_1 m_2/(m_1+m_2)$ is the inertia moment of the molecule. Thus,

$$\begin{aligned}
\mathcal{Z} &= e^{-\beta\varepsilon_0} \cdot \mathcal{Z}_{\text{rot}} \cdot \mathcal{Z}_{\text{vib}} \\
\mathcal{Z}_{\text{rot}} &= \sum_{K=0}^{\infty} (2K+1)\, e^{-\beta\hbar^2 K(K+1)/2I} , \\
\mathcal{Z}_{\text{vib}} &= \sum_{v=0}^{\infty} e^{-\beta\hbar\omega(v+1/2)} .
\end{aligned}$$

As a result,

$$F = -NT \log \left[\frac{eV}{N} \left(\frac{mT}{2\pi\hbar^2} \right)^{3/2} \right] + N\varepsilon_0 + F_{\text{rot}} + F_{\text{vib}} .$$

The 2 last terms are defined according to the definition $\delta F_\alpha = -NT \log \mathcal{Z}_\alpha$. Introducing the translational contributions to the specific heat and chemical constant as

$$\begin{aligned}
\zeta_c &= \zeta_{\text{tr}} + \zeta_{\text{rot}} + \zeta_{\text{vib}} , \quad \zeta_{\text{tr}} = (3/2) \log(m/2\pi\hbar^2) , \\
c_V &= c_{\text{tr}} + c_{\text{rot}} + c_{\text{vib}} , \quad c_{\text{tr}} = 3/2 , \\
c_P &= c_{\text{tr}} + c_{\text{rot}} + c_{\text{vib}} + 1 .
\end{aligned}$$

As an example let us consider the case of large temperatures, when $T \gg \hbar^2/2I$. Then one can replace summation by the integration to obtain

$$\mathcal{Z}_{\text{rot}} = \int_0^\infty (2K+1)\, e^{-\beta\hbar^2 K(K+1)/2} = 2TI/\hbar^2 , \quad c_{\text{rot}} = 1 , \quad \zeta_{\text{rot}} = \log(2I/\hbar^2) .$$

In this case

$$c_V = 5/2 , \quad c_P = 7/2 .$$

In the opposite limit it is sufficient to retain the first two terms,

$$\mathcal{Z}_{\text{rot}} = 1 - 3e^{-\beta\hbar^2/I}$$

to obtain

$$F_{\text{rot}} = -3NTe^{-\beta\hbar^2/I} , \quad c_{\text{rot}} = 3(\hbar^2/IT)^2 3e^{-\hbar^2/IT} .$$

Note that above expressions are valid only for the molecules with the *different* atoms.

If the atoms are equivalent, then at large temperatures two opposite directions correspond to the same physical state. As a result, the rotational partition function must be halved. The quantum case is more difficult and we do not discuss it.

The *vibrational* degrees of freedom can be easily accounted by direct summation of geometric series. Using the lowest vibrational level as the origin for the energies, we obtain

$$\begin{aligned}
\mathcal{Z}_{\text{vib}} &= \left[1-e^{-\beta\hbar\omega}\right], \\
F_{\text{vib}} &= NT\log\left[1-e^{-\beta\hbar\omega}\right], \\
S_{\text{vib}} &= -N\log\left[1-e^{-\beta\hbar\omega}\right]+N\frac{\beta\hbar\omega}{e^{\beta\hbar\omega}-1}, \\
U_{\text{vib}} &= N\frac{\hbar\omega}{e^{\beta\hbar\omega}-1}, \\
c_{\text{vib}} &= \frac{(\beta\hbar\omega)^2 e^{\beta\hbar\omega}}{\left(e^{\beta\hbar\omega}-1\right)^2}.
\end{aligned}$$

The graphs for the rotational and vibrational contributions to the specific heat are shown in Fig. 6.1.

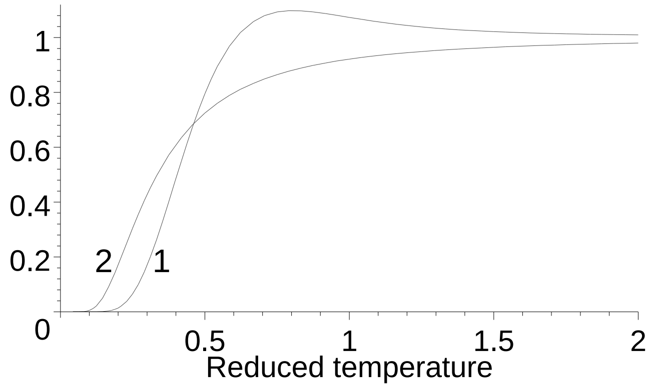

Figure 6.1: Contributions to the specific heat from rotational (1) and vibrational (2) degrees of freedom. Temperature is measured in units of $\hbar^2/2I$ in the graph (1) and in units of $\hbar\omega$ in the graph (2).

Exercise: Calculate approximate expressions for the case of low and high temperatures and discuss their physical meaning.

Quantum Gases

Statistics of identical particles

Consider N indistinguishable particles moving in a common potential,

$$\hat{\mathcal{H}} = \sum_{i=1}^{N} \hat{\mathcal{H}}_i , \quad \hat{\mathcal{H}}_i = \frac{\hat{\mathbf{p}}_i^2}{2m} + V(\hat{\mathbf{r}}_i) .$$

Here $\hat{\mathbf{p}}_i = -i\hbar\nabla_i$ and $\hat{\mathbf{r}}_i$ ore operators for the momentum and coordinate of the i-th particle. Let us denote single-particle eigenenergies as ε_k and the states as ψ_k,

$$\hat{\mathcal{H}}_i \psi_k(\mathbf{r}_i) = \varepsilon_k \psi_k(\mathbf{r}_i) .$$

Specifying the total wave function as $\Psi_s = \prod_i \psi_{k_i}(\mathbf{r}_i)$ we obtain $E_s = \sum_i \varepsilon_{k_i}$ since the Hamiltonian is the sum of independent terms.

Now we have to use the concept that the particles are indistinguishable. Introducing the permutation operator $\hat{P}$ which interchanges i-th and j-th particle, $\mathbf{r}_i \leftrightarrow \mathbf{r}_j$. Then, one has to require

$$\left[\hat{\mathcal{H}}, \hat{P}\right] = 0 , \quad \text{or} \quad \hat{P}\hat{\mathcal{H}}\hat{P}^{-1} = \hat{\mathcal{H}} .$$

According to the quantum theory, the proper many-body wave function must be an eigenstate of the permutation operator,

$$\hat{P}\Psi_s = p\Psi_s .$$

Since double permutation is just the identity operator, $\hat{P}^2 = \hat{I}$, we get

$$p^2 = 1 \quad \to \quad p = \pm 1 .$$

The particles with $p = 1$ called the *bosons*, while the particles with $p = -1$ are called the *fermions*.

The antisymmetric wave function vanishes if two particles are in the same state. This property is called the Pauli principle.

Example: Particles in a Harmonic Potential: Ignoring zero-point vibrations we have $\varepsilon_n = n\hbar\omega$, $n = 0, 1, 2 \ldots$. Then

$$Z_1 = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x} , \quad x \equiv e^{-\beta\hbar\omega} .$$

For the Maxwell-Boltzmann (MB) statistics we have

$$Z_2^{MB} = \frac{Z_1^2}{2!} = \frac{1}{2} \frac{1}{(1-x)^2} .$$

Note: The product Z_1^2 contains 1 configuration where 2 particles occupy the same state. By dividing this product by 2! we ascribe the weight 1/2 to this configuration. This procedure is wrong from the point of view of quantum statistics. Indeed, bosons are allowed to occupy the same single-particle state, so the the diagonal term

$$D=\sum_{n=0}^{\infty}x^{2n}=\frac{1}{1-x^2}$$

must be added with the weight 1/2.

$$Z_2^{BE}=Z_2^{MB}+\frac{1}{2}D\,, \quad \to \quad Z_2^{BE}=\frac{1}{(1-x)(1-x^2)}\,.$$

In a similar way, the diagonal term must be subtracted for Fermi-Dirac (FD)statistics,

$$Z_2^{FD}=Z_2^{MB}-\frac{1}{2}D\,, \quad \to \quad Z_2^{FD}=\frac{x}{(1-x)(1-x^2)}\,.$$

The difference is negligible for $1-x\ll 1$ (large temperature).

The Fermi-Dirac Statistics

This statistics is based on the *Pauli principle*:

> Each quantum state cannot be simultaneously occupied by more than 1 particle.
>
> The particles obeying the Fermi-Dirac statistics are called the *fermions*.

Derivation from a grand canonical ensemble

Consider the system of particles with single-particle levels ε_i and degeneracy factors g_i. If the state is occupied, then there are only g_i-1 states at the i-th level for the second particles, the number of available states for two particles being $g_i(g_1-1)/2$. The factor $1/2$ appears because the particles are non-distinguishable. For n_i particles we have

$$W=\prod_i\frac{g_i!}{n_i!(g_i-n_i)!}\,.$$

Now we apply the conventional approach. Namely, consider a transition $n_i\to n_i+1$. The energy change is ε_i, the change in the particle number is $\Delta N=1$, while the entropy change is (prove!)

$$\Delta S=\log\frac{g_i-n_i}{n_i+1}\approx\log\frac{g_i-n_i}{n_i}\,.$$

Since

$$T\,\Delta S=\Delta U-\mu\Delta N=\varepsilon_i-\mu$$

we obtain

$$\frac{g_i}{n_i}-=e^{\beta(\varepsilon_i-\mu)} \quad \to n_i=\frac{g_i}{e^{\beta(\varepsilon_i-\mu)}+1}\,.$$

As a result, the probability for a state i to be occupied is

$$f_i = \frac{1}{e^{\beta(\varepsilon_i-\mu)}+1}\,, \tag{6.10}$$

the normalization condition (which determines μ) being

$$N = \sum_i g_i f_i = \sum_i \frac{g_i}{e^{\beta(\varepsilon_i-\mu)}+1}\,.$$

Derivation from the Gibbs distribution

Consider the particles belonging to i-th state. We have the general expression

$$\Omega_i = -T\,\log\sum_{n_i} e^{\beta(n_i\mu - n_i\varepsilon_i)} = -T\,\log\sum_{n_i}\left(e^{\beta(\mu-\varepsilon_i)}\right)^{n_i}. \tag{6.11}$$

Now, since $n_i = 0$ or 1, we get

$$\Omega_i = -T\,\log\left(1+e^{(\beta\mu-\varepsilon_i)}\right). \tag{6.12}$$

The following way is standard:

$$f_i = \bar{n_i} = -\frac{\partial\Omega_i}{\partial\mu} = \frac{1}{e^{\beta(\varepsilon_i-\mu)}+1}\,. \tag{6.13}$$

The total Landau free energy is

$$\Omega = -T\,\log\sum_i\left(1+e^{\beta(\mu-\varepsilon_i)}\right). \tag{6.14}$$

The Bose-Einstein distribution

For *bosons* the number of particles occupying any state can be *arbitrary*. A proper picture is g_i boxes separated by g_i-1 walls. To count the number of configurations one can count the different grouping of g_i-1 walls and n_i bosons. As a result,

$$W = \prod_i \frac{(n_i+g_i-1)!}{n_i!(g_i-1)!} \quad \to \Delta S \approx \log\frac{g_i+n_i}{n_i}\,.$$

Consequently,

$$f_i = \frac{n_i}{g_i} = \frac{1}{e^{\beta(\varepsilon_i-\mu)}-1}\,. \tag{6.15}$$

We can derive this expression also from the general formula (6.11) by performing summation over n_i from 0 to ∞. Using the properties of geometric series we obtain,

$$\Omega_i = T\ln\left(1-e^{\beta(\mu-\varepsilon_i)}\right). \tag{6.16}$$

Again, we recover the expression (6.15) as $f_i = -\partial\Omega_i/\partial\mu$. It is worthwhile noting that Eq. (6.15) is applicable only if

$$e^{\beta(\mu-\varepsilon_i)} < 1 \quad \to \quad \mu < 0\,.$$

We immediately recover the Boltzmann distribution at $n_i \ll g_i$ by replacing $(n_i+g_i-1)!$ by $g_i^{n_i}$.

6.2 Ideal gases out of equilibrium

Let us assume that we label the states of the system by the subscript i, the degeneracy factor and the occupation number of ith state being g_i and n_i, respectively. Then the set $\{n_i\}$ describes the state of the system if it is also out of equilibrium. Since the particles are *independent*,

$$W = \prod_i \Delta\Gamma_i$$

For the Boltzmann gas one can place *each* of n_i particles in one of the possible g_i state. Then the number of different physical configurations is

$$\Delta\Gamma_i = g_i^{n_i}/n_i!\,. \tag{6.17}$$

As a result,

$$S = \ln W = \sum \ln(g_i^{n_i}/n_i!) = \sum (n_i \ln g_i - \ln n_i!)\ . \tag{6.18}$$

Now we can use the Stirling's formula to get

$$S = \sum_i n_i \ln(e g_i/n_i) \equiv \sum_i g_i f_i \ln(e/f_i)\,. \tag{6.19}$$

Here we have introduced the distribution function $f_i \equiv n_i/g_i$ as the probability for ith state to be occupied.

Let us now require that the entropy must be a maximum for the equilibrium state. Then the problem is to find f_i such that the sum (6.19) has the maximum value possible under the restrictions

$$\sum_i n_i = \sum_i g_i f_i = N \quad \text{(particle number conservation)}\,, \tag{6.20}$$

$$\sum_i \varepsilon_i n_i = \sum_i g_i \varepsilon_i f_i = E \quad \text{energy conservation)}\,. \tag{6.21}$$

Requiring that

$$\partial(S + \mu\beta N - \beta E)/\partial f_i = 0\,,$$

one obtains

$$g_i(-\ln f_i + \alpha + \beta\varepsilon_i) = 0 \quad \to \quad f_i = e^{\beta(\mu-\varepsilon_i)}\,. \tag{6.22}$$

For the Fermi gas, *not more than one* particle can be in each state. Thus one has to count the number of way to fill n_i states from g_i available ones, and $\Delta\Gamma_i = g_i!/n_i!(g_i - n_i)!$. Thus

$$S = -\sum_i g_i \left[f_i \ln f_i + (1 - f_i)\ln(1 - f_i)\right]\ . \tag{6.23}$$

Again (Check!), requiring this expression to have a maximum under conditions (6.20) and (6.21) we arrive at the Fermi-Dirac distribution (6.13).

For the Bose-Einstein statistics one has to calculate the number of ways of distributing n_i particles among g_i states. This problem can be mapped on the problem of distributing n_i identical balls among g_i urns. Let us imagine that the latter are separated by by $g_i - 1$ vertical strokes placed at intervals along the points. For example, the diagram

.|...||....|..

represents 10 balls among 5 urns as $\{1,3,0,4,2\}$. The total number of places (occupied both by strokes and points) is then $(g_i + n_1 - 1)$. The number of configuration is just the number of ways to choose $g_1 - 1$ positions for the strokes. Thus, $\Delta\Gamma_i = ((g_i + n_1 - 1)!/n_i!(g_i - 1)!$, and

$$S = \sum_i g_i \left[(1 + f_1)\ln(1 + f_i) - f_i \ln f_i\right] . \tag{6.24}$$

The expressions (6.23) and (6.24) tend to the Boltzmann formula (6.19) at $n_i \ll g_i$ since

$$g_i! \approx (g_i - n_i)! g_i^{n_i} , \; (g_i + n_i - 1)! \approx (g_i - 1)! g_i^{n_i} \quad \text{at} \quad n_i \ll g_i .$$

Finally, let us specify the expression for the entropy of Bose gas at $f_i \gg 1$ since we will need this expression later to discuss the Bose condensation. We have $\Delta\Gamma_i \approx n_i^{(g_i - 1)}/(g_i - 1)!$ and

$$S = \sum_i g_i \ln(e f_i) . \tag{6.25}$$

6.3 Fermi and Bose gases of elementary particles

For elementary particles we have

$$\varepsilon = \frac{\mathbf{p}^2}{2m} , \quad g = (2s + 1) .$$

Here g is the spin degeneracy. Thus the equation for the chemical potential has the form, see Eq. (6.2),

$$\frac{N}{V} = \int d\varepsilon D(\varepsilon) f(\varepsilon) = g D_0 \int_0^\infty \frac{\sqrt{\varepsilon}\, d\varepsilon}{e^{\beta(\varepsilon - \mu)} \pm 1} = D_0 T^{3/2} \int_0^\infty \frac{\sqrt{z}\, dz}{e^{z - \beta\mu} \pm 1} .$$

Here "+" stands for fermions while "-" stands for bosons. In a similar way,

$$\Omega = \mp V D_0 T \int_0^\infty d\varepsilon \sqrt{\varepsilon} \log\left(1 \pm e^{\beta(\mu - \varepsilon)}\right) = -\frac{2}{3} V D_0 \int_0^\infty d\varepsilon\, \varepsilon^{3/2} f(\varepsilon) .$$

The second inequality is obtained using integration by parts. We recognize that

$$\Omega = -(2/3) \sum_i g_i \varepsilon_i f_i = -(2/3) \int_0^\infty D(\varepsilon) \varepsilon f(\varepsilon)\, d\varepsilon = -(2/3) U . \tag{6.26}$$

Since $\Omega = -PV$ we arrive at the following relationship:

$$PV = \frac{2}{3} U .$$

This result holds also for Boltzmann distribution.

Using the substitution of the variables $z = \beta\varepsilon$ we observe that

$$\Omega = -PV = VT^{5/2}\eta(\mu/T)\,,$$

where η is a function of a single variable. Since

$$\frac{S}{V} = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} \quad \text{and} \quad \frac{N}{V} = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V}$$

the ratio S/N depends only on the ratio μ/T.

If the process is adiabatic, then $S = \text{const}$, and μ/T must be constant. Consequently,

$$VT^{3/2} = \text{const}\,, \quad PV^{5/3} = \text{const}\,, \quad T^{5/2}P = \text{const}.$$

The exponents, however, have nothing to do with the ratio of specific heats, c_V/c_P, since the conditions $c_P/c_V = 5/3$ and $c_P - c_V = 1$ do not hold any more.

Equation of state

The equation of state can be written in a parametric form:

$$\begin{aligned}\frac{P(V,T)}{D_0T^{5/2}} &= \int_0^\infty \frac{z^{3/2}\,dz}{e^{z-\beta\mu(T,V)} \pm 1}\,,\\ \frac{N(V,T)}{VD_0T^{3/2}} &= \int_0^\infty \frac{z^{1/2}\,dz}{e^{z-\beta\mu(T,V)} \pm 1}\,.\end{aligned}$$

We have to know some tips to calculate the involved integrals.

In the limiting case of the Boltzmann gas, $e^{\beta\mu} \ll 1$, one can expand

$$\frac{1}{e^{z-\beta\mu} \pm 1} \approx e^{\beta\mu - z}\left(1 \mp e^{\beta\mu - z} \pm \ldots\right)\,.$$

The lowest approximation leads to the Boltzmann value of the chemical potential and to Clapeyron's equation $PV = NT$ (Check!). The next term provides proper corrections. From Eq. (6.26) we obtain

$$\Omega \approx \Omega_{\text{Bol}}\left(1 \mp \frac{1}{2^{5/2}}e^{\beta\mu}\right) = \approx \Omega_{\text{Bol}} \pm \frac{gVm^{3/2}T^{5/2}}{16\pi^{3/2}\hbar^3}e^{2\beta\mu}$$

where Ω_{Bol} is the Boltzmann value for the Landau free energy. Substituting the Boltzmann value for the chemical potential in the second term and using the concept of *small increments* (4.20) one can express the free energy as

$$F = F_{\text{Bol}} \pm \frac{\pi^{3/2}}{2g} \cdot \frac{N^2\hbar^3}{VT^{1/2}m^{3/2}}\,.$$

Finally, differentiating this expression with respect to volume we arrive at the equation of state:

$$PV = NT\left[1 \pm \frac{\pi^{3/2}}{2g}\frac{N\hbar^3}{V(mT)^{3/2}}\right]. \tag{6.27}$$

The conditions for the correction to be shall is naturally the same as for applicability of the Boltzmann statistics. Quantum effects behave and an additional *attraction* for bosons and as an additional *repulsion* for fermions.

A degenerate electron gas

For electron spin equals to 1/2, and $g = 2s+1 = 2$. At zero temperature, one can approximate the Fermi function

$$f(\varepsilon) = \frac{1}{e^{\beta(\varepsilon-\mu)}+1} \approx \left\{\begin{array}{lll} 1 & \text{at} & \varepsilon < \mu \\ 0 & \text{at} & \varepsilon > \mu \end{array}\right.$$

Thus there is a maximum energy

$$\mu|_{T=0} \equiv \varepsilon_F = p_F^2/2m\,.$$

Since

$$2\cdot\frac{4\pi p^2\,dp\cdot V}{(2\pi\hbar)^3} = V\frac{p^2\,dp}{\pi^2\hbar^3}$$

we have

$$\frac{N}{V} = \frac{p_F^3}{3\pi^2\hbar^3}, \quad \to \quad p_F = (3\pi^2)^{1/3}(N/V)^{1/3}\hbar, \quad \varepsilon_F = (3\pi^2)^{2/3}(N/V)^{2/3}(\hbar^2/2m)\,.$$

Limiting the integration of the absolute value of the electron moments by p_F we obtain

$$\begin{aligned} \frac{U}{V} &= \frac{p_F^5}{10m\pi^2\hbar^3} = \frac{3(3\pi^2)^{2/3}}{10}\frac{\hbar^2}{m}\left(\frac{N}{V}\right)^{5/3}, \\ P &= \frac{(3\pi^2)^{2/3}}{5}\frac{\hbar^2}{m}\left(\frac{N}{V}\right)^{5/3}. \end{aligned}$$

To calculate specific heat at low temperatures one has to be able to find temperature-dependent corrections. Let us demonstrate the way to calculate these corrections. Consider

$$I = \int_0^\infty \frac{f(\varepsilon)\,d\varepsilon}{e^{\beta(\varepsilon-\mu)}+1},$$

where $f(\varepsilon)$ is an arbitrary function such that the integral converges. Introducing $z = \beta(\varepsilon-\mu)$ and remembering that $\beta \equiv 1/T$ we get

$$I = T\int_{-\beta\mu}^{\infty}\frac{f(\mu+Tz)}{e^z+1}\,dz = T\int_0^{\beta\mu}\frac{f(\mu-Tz)}{e^{-z}+1}\,dz + T\int_0^\infty\frac{f(\mu+Tz)}{e^z+1}\,dz\,.$$

Now we can replace in the 1st integral

$$\frac{1}{e^{-z}+1}=1-\frac{1}{e^{z}+1}$$

to obtain

$$I=T\int_0^{\beta\mu}f(\mu-Tz)\,dz+T\int_0^{\infty}\frac{f(\mu+Tz)-f(\mu-Tz)}{e^z+1}\,dz+\int_{\beta\mu}^{\infty}\frac{f(\mu-Tz)}{e^z+1}\,dz\,.$$

The last term is exponentially small since $\beta\mu\gg 1$. Returning to the initial variables in the first term, we obtain

$$I=\int_0^{\mu}f(\varepsilon)\,d\varepsilon+T\int_0^{\infty}\frac{f(\mu+Tz)-f(\mu-Tz)}{e^z+1}\,dz\,.$$

At low temperature one can expand the second integrand as

$$f(\mu+Tz)-f(\mu-Tz)=2Tf'(\mu)z+\frac{2}{3!}T^3f'''(\mu)z^3+\cdots\,.$$

As a result,

$$\begin{aligned} I &= \int_0^{\mu}f(\varepsilon)\,d\varepsilon+2T^2f'(\mu)\int_0^{\infty}\frac{z\,dz}{e^z+1}+\frac{1}{3}T^4f'''(\mu)\int_0^{\infty}\frac{z^3\,dz}{e^z+1}+\cdots \\ &= \int_0^{\mu}f(\varepsilon)\,d\varepsilon+\frac{\pi^2}{6}T^2f'(\mu)+\frac{7\pi^4}{360}T^4f'''(\mu)+\cdots\,. \end{aligned} \tag{6.28}$$

Now we can calculate temperature-dependent corrections to the thermodynamic quantities. Substituting $f(\varepsilon)=\varepsilon^{3/2}$ we get

$$\Omega=\Omega_0-VT^2\frac{m^{3/2}\sqrt{2\mu}}{6\hbar^3}\,.$$

Since the second term is a small correction we can replace $\mu\rightarrow\varepsilon_F$ in that term. Using the theorem of small increments we find

$$\begin{aligned} F &= F_0-\frac{1}{2}\gamma NT^2\left(\frac{V}{N}\right)^{2/3}, \quad \gamma\equiv\left(\frac{\pi}{3}\right)^{2/3}\frac{m}{\hbar^2}\,, \\ S &= \gamma NT\left(\frac{V}{N}\right)^{2/3}, \\ C &= \gamma NT\left(\frac{V}{N}\right)^{2/3}\sim\frac{T}{\varepsilon_F}\,. \end{aligned}$$

A degenerate Bose gas

The chemical potential of bosons is negative. It is given by the equation

$$\frac{N}{V}=\frac{g(mT)^{3/2}}{2^{1/2}\pi^2\hbar^3}I(\beta|\mu|)\,,\quad I(z)=\int_0^{\infty}\frac{\sqrt{x}\,dx}{e^{z+x}-1}\,. \tag{6.29}$$

It increases with the temperature decrease and reaches zero at the temperature T_0 given by the equation

$$\frac{N}{V} = \frac{g(mT_0)^{3/2}}{2^{1/2}\pi^2\hbar^3} I(0)\,.$$

Since $I(0) = (\sqrt{\pi}/2)\zeta(3/2)$ where $\zeta(x)$ is the Riemann's ζ-function,

$$T_0 \approx \frac{3.31}{g^{2/3}}\frac{\hbar^2}{m}\left(\frac{N}{V}\right)^{2/3}.$$

From Eq. (6.29) we obtain

$$(\beta T_0)^{3/2} = I(\beta|\mu|)/I(0)\,. \tag{6.30}$$

The ratio $\mathcal{F}(z) \equiv I(z)/I(0)$ is plotted in Fig. 6.2. The dependence of the product $\beta|\mu|$ versus βT_0 can be easily found graphically from this plot.

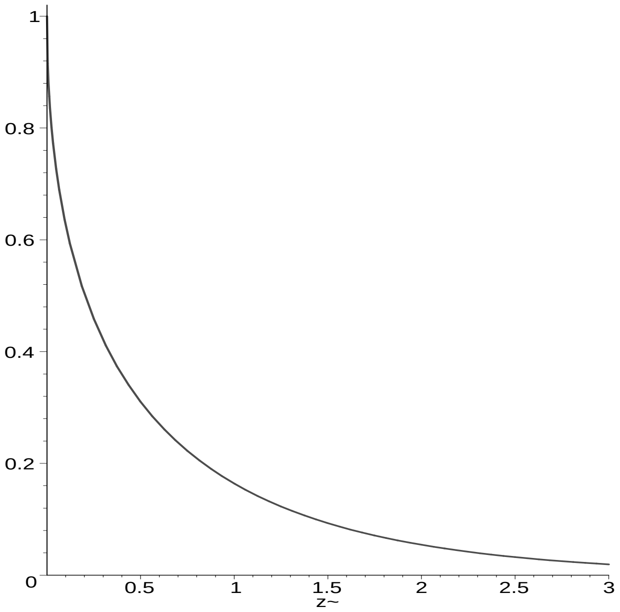

Figure 6.2: Plot of the ratio $\mathcal{F}(z) \equiv I(z)/I(0)$. Chemical potential can be found by graphical solution of Eq. (6.30).

What happens at $T < T_0$? The above equation is based on replacement of the summation over discrete states by the integration. This is not true for bosons because they can *condense* in the ground state with $\varepsilon_k = 0$. We omit this term in our integration because $\sqrt{\varepsilon_k} = 0$ for the ground state.

As a result, at $T < T_0$ the particles with $\varepsilon > 0$ are distributed according to the BE distribution with $\mu = 0$,

$$dn_\varepsilon = \frac{gm^{3/2}V}{2^{1/2}\pi^2\hbar^3}\frac{\sqrt{\varepsilon}\,d\varepsilon}{e^{\beta\varepsilon}-1}\,.$$

The total number of such particles is

$$N_{\varepsilon>0} = \frac{gm^{3/2}V}{2^{1/2}\pi^2\hbar^3}\int_0^\infty = \frac{\sqrt{\varepsilon}\,d\varepsilon}{e^{\beta\varepsilon}-1} = N\left(\frac{T}{T_0}\right)^{3/2}.$$

The remaining

$$N_{\varepsilon=0} = N\left[1-\left(\frac{T}{T_0}\right)^{3/2}\right]$$

particles are in the lowest state with zero energy. The energy and other thermodynamic quantities can be obtained from general expression for Bose gas by putting $\mu = 0$. In this way we obtain (check!)

$$\begin{aligned} U &= 0.770NT(T/T_0)^{3/2}, \\ C_V &= 5U/3T, \\ F &= U-TS = -(2/3)U = \Omega \quad (since)\ \mu = 0, \\ P &= -(\partial F/\partial V)_T = 0.0851gm^{3/2}T^{5/2}/\hbar^3. \end{aligned} \tag{6.31}$$

Thus at $T \leq T_0$ the pressure is $\propto T^{5/2}$ and it is *volume-independent*. The reason is that the particles in the condensate do not contribute to the pressure.

6.4 Black-body radiation

Electromagnetic radiation which in a thermal equilibrium is called the *black-body radiation*. It can be expressed as gas of non-interacting photons. To achieve the equilibrium, photons must interact with some amount of matter which is able to absorb and emit photons. Thus we arrive at the situation with variable number of photons, and the chemical potential μ should be fixed. Since the free energy must be minimum

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad \text{and} \quad \frac{\partial F}{\partial N} = 0$$

we obtain

$$\mu = 0.$$

The photons are specified by the wave vector $\mathbf{k}$ (the momentum being $\hbar\mathbf{k}$) and frequencies $\omega_{\mathbf{k}} = ck$ (the energy being $\hbar\omega_{\mathbf{k}} = \hbar ck$), as well as by polarization. Since photons are *transverse*, there are 2 directions of polarization, and the one obtains for the density of states

$$2\cdot V\frac{d^3k}{(2\pi)^3} = 2\cdot V\frac{4\pi k^2\,k}{(2\pi)^3} = V\frac{\omega^2\,d\omega}{\pi^2c^3} \equiv VD(\omega)\,d\omega.$$

Here $D(\omega) = \omega^2/\pi^2c^3$. In this way we obtain

$$dn_\omega = V\frac{\omega^2}{\pi^2c^3}N(\omega)\,d\omega$$

where

$$N(\omega) \equiv \frac{1}{e^{\beta\hbar\omega}-1}$$

is called the Planck distribution. The energy distribution is obtained by multiplication by $\hbar\omega$ (Planck, 1900),

$$dE_\omega = V\frac{\hbar\omega^3}{\pi^2c^3}\frac{1}{e^{\beta\hbar\omega}-1}\,d\omega\,,$$

its low-frequency asymptotics is the *Rayleigh-Jeans law*,

$$dE_\omega = V\frac{T\,\omega^2}{\pi^2c^3}\,d\omega\,,$$

while the high frequency one,

$$dE_\omega = V\frac{\hbar\omega^3}{\pi^2c^3}e^{-\beta\hbar\omega}\,d\omega\,,$$

is the *Wien's* law.

The thermodynamic quantities can be derived in a usual way, e. g.,

$$F = T\int_0^\infty D(\omega)\,\log\left(1-e^{-\beta\hbar\omega}\right) = -4\sigma VT^4/3c\,.$$

Here $\sigma \equiv \pi^2/60\hbar^3c^3$ is the Stefan-Boltzmann constant. It temperature is measured in degrees, then

$$\sigma \equiv \pi^2k^4/60\hbar^3c^3 = 5.67\times10^{-5}\ \text{g/sec}^3\text{deg}^4\,.$$

For other quantities we obtain,

$$S = 16\sigma T^3/3c\,,\ U = 4\sigma VT^4/c = -3F\,,\ C_V = 16\sigma VT^3/c\,,\ PV = U/3 = 4\sigma T^4/c\,.$$

A body in thermal equilibrium with black-body radiation: Let us denote

$$\mathcal{E}_\omega = \frac{1}{4\pi V}\frac{dE_\omega}{d\omega} = \frac{\hbar\omega^3}{4\pi^3c^2\left(e^{\beta\hbar\omega}-1\right)}$$

for the spectral density of black-body radiation per unit volume and unit solid angle. Then the energy flux density leaving each point of the body is $c\mathcal{E}_\omega\,do\,d\omega$. The *incident* energy flux on unit area of the surface at an angle θ to the normal is therefore $c\mathcal{E}_\omega\cos\theta\,do$, $do = 2\pi\sin\theta\,d\theta$. Let us not define the *absorbing power*, i. e. fraction of the incident energy absorbed by the body, as $A(\omega,\theta)$. Then the energy absorbed per unit time and surface area is

$$c\mathcal{E}_\omega A(\omega,\theta)\cos\theta\,do\,d\omega\,.$$

For simplicity, we can also assume that the body does not *scatter* radiation and is not fluorescent, i. e. that reflection occurs without change in the angle θ and in frequency ω. If we also assume that the radiation does no *pass through the body* we arrive at the balance equation between the radiation density and the intensity of emission $J(\omega,\theta)\,do\,d\omega$ in the same direction:

$$\frac{J(\omega,\theta)}{A(\omega,\theta)} = c\mathcal{E}_\omega\cos\theta \quad \to \quad \text{a universal function of } \omega \text{ and } \theta\,.$$

This is the *Kirchhof's law*.[1] If $A(\omega,\theta) = 1$, i. e. if the body absorbs all the radiation, it is called the *black body*. Then

$$J(\omega,\theta) = c\mathcal{E}_\omega \cos\theta\,, \quad \to \quad J_0 = \int do\, d\omega\, J(\omega,\theta) = cU/4V = \sigma T^4\,.$$

6.5 Lattice Vibrations. Phonons

For small displacement on an atom from its equilibrium position one can expand the potential energy near its minimal value

$$\begin{aligned} V(R) &= V(R_0) + \left(\frac{dV}{dR}\right)_{R_0} (R-R_0) + \frac{1}{2}\left(\frac{d^2V}{dR^2}\right)_{R_0} (R-R_0)^2 + + \\ &+ \frac{1}{6}\left(\frac{d^3V}{dR^3}\right)_{R_0} (R-R_0)^3 + \cdots \end{aligned} \tag{6.32}$$

If we expand the energy near the equilibrium point and denote

$$\left(\frac{d^2V}{dR^2}\right)_{R_0} \equiv C > 0, \quad \left(\frac{d^3V}{dR^3}\right)_{R_0} \equiv -2\gamma > 0$$

we get the following expression for the restoring force for a given displacement $x \equiv R - R_0$

$$F = -\frac{dV}{dx} = -Cx + \gamma x^2 \tag{6.33}$$

The force under the limit $F = -Cx$ is called *quasi elastic*.

One-Atomic Linear Chain

Dispersion relation

We start with the simplest case of one-atomic linear chain with nearest neighbor interaction. If one expands the energy near the equilibrium point for the nth atom and use quasi elastic approximation (6.33) he comes to the Newton equation (see Fig. /reff-ph1)

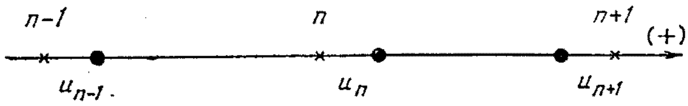

Figure 6.3:

$$m\ddot{u}_n + C(2u_n - u_{n-1} - u_{n+1}) = 0. \tag{6.34}$$

---

[1] If there is scattering, then the Kirchhof's law can be formulated only for integrals over frequencies and angles.

To solve this *infinite* set of equations let us take into account that the equation does not change if we shift the system *as a whole* by the quantity a times an integer. We can fulfill this condition automatically by searching the solution as

$$u_n = Ae^{i(qan-\omega t)}. \tag{6.35}$$

Immediately we get

$$\omega = \omega_m \left| \sin \frac{qa}{2} \right|, \quad \omega_m = 2\sqrt{C/m}. \tag{6.36}$$

The expression (6.36) is called the *dispersion law*. It differs from the dispersion relation for an homogeneous string, $\omega = sq$. The spectrum is shown in Fig. 6.4. Another important feature is that if we replace the wave number q as

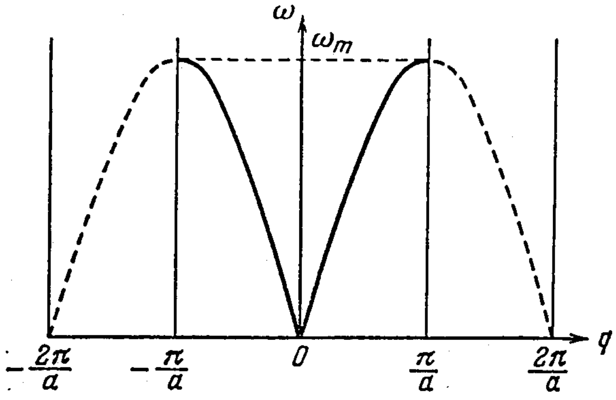

Figure 6.4: Dispersion law for vibrations in a one-atomic chain. Periodic character is clearly seen.

$$q \to q' = q + \frac{2\pi g}{a},$$

where g is an integer, the solution (6.35) does not change (because $\exp(2\pi i \times \text{integer}) = 1$). Consequently, it is impossible to discriminate between q and q' and it is natural to choose the region

$$-\frac{\pi}{a} \le q \le \frac{\pi}{a} \tag{6.37}$$

to represent the dispersion law in the whole q-space. Note that there is the *maximal frequency* ω_m that corresponds to the *minimal* wave length $\lambda_{\min} = 2\pi/q_{\max} = 2a$. The maximal frequency is a typical feature of discrete systems vibrations.

Now we should recall that any crystal is *finite* and the translation symmetry we have used fails. The usual way to overcome the problem is to take into account that actual number L of sites is large and to introduce *Born-von Karmann* cyclic boundary conditions

$$u_{n\pm L} = n_n. \tag{6.38}$$

This condition make a sort of ring of a very big radius that physically does not differ from the long chain.[2] Immediately, we get that the wave number q should be *discrete*. Indeed, substituting the condition (6.38) into the solution (6.35) we get $\exp(\pm iqaL) = 1,\ qaL = 2\pi g$ with an integer g. Consequently,

$$q = \frac{2\pi}{a}\frac{g}{L}, \quad -\frac{L}{2} < g < \frac{L}{2} \tag{6.39}$$

(it is convenient to consider L as a large even number). So, for a linear chain, the wave number q takes L discrete values in the interval $(-\pi/a, \pi/a)$. Note that this interval is just the same as the Wigner-Zeitz cell of the one-dimensional reciprocal lattice.

Density of States

Because of the discrete character of the vibration states one can calculate *the number of states*, $D(\omega)$, with different q in the frequency interval $\omega, \omega + d\omega$. One easily obtains (check!)

$$D(\omega) = \frac{2L}{\pi}\frac{1}{\sqrt{\omega_m^2 - \omega^2}}. \tag{6.40}$$

It is divergent at $\omega \to \omega_m$ which is a feature of 1D systems. The density of states $D(\omega)$ is shown in Fig. 6.5.

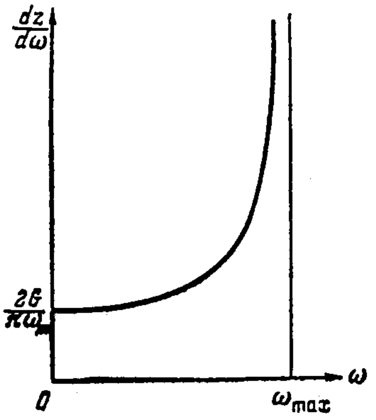

Figure 6.5: Density of vibrational states in a one-atomic chain. There is a singularity at the maximal frequency that is a typical feature of 1D systems.

Phase and Group Velocity

Now we discuss the properties of long wave vibrations. At small q we get from Eq. (6.36)

$$\omega = sq, \tag{6.41}$$

---

[2] Note that for small structures of modern electronics this assumption need revision. Violation of this assumption leads to the specific interface modes.

where

$$s = a\sqrt{\frac{C}{m}} \tag{6.42}$$

is the sound velocity in a homogeneous elastic medium. In a general case, the sound velocity becomes q-dependent, i. e. there is the dispersion of the waves. One can discriminate between the phase (s_p) and group (s_g) velocities. The first is responsible for the propagation of the equal phase planes while the last one describes the energy transfer. We have

$$\begin{aligned} s_p &= \frac{\omega}{|q|} = s\left|\frac{\sin(aq/2)}{aq/2}\right| , \\ s_g &= \left|\frac{d\omega}{dq}\right| = s|\cos(aq/2)| . \end{aligned} \tag{6.43}$$

At the boundaries of the interval we get $s_p = (2/\pi)s$ while $s_g = 0$ (boundary modes cannot transfer energy).

Diatomic Chain. Acoustic and Optical branches.

We use this case to discuss vibrations of compound lattices. The corresponding picture is shown in Fig. 6.6

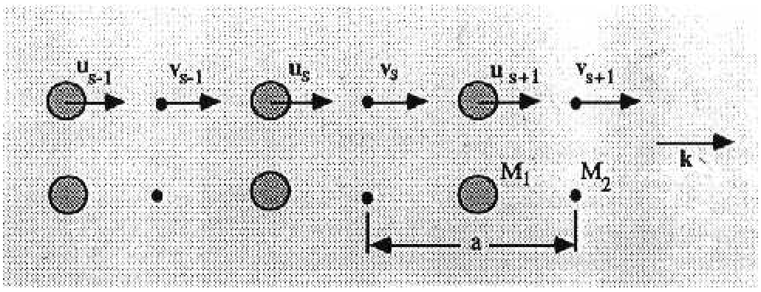

Figure 6.6: On a two-atomic chain.

One can see that the elementary cell contains 2 atoms. If we assume the elastic constants to be $C_{1,2}$ we come to the following equations of motion:

$$\begin{aligned} m_1\ddot{u}_n &= -C_1(u_n - v_n) - C_2(u_n - v_{n-1}) , \\ m_2\ddot{v}_n &= -C_1(v_n - u_n) - C_2(v_n - u_{n-1}) . \end{aligned} \tag{6.44}$$

It is natural to use once more the translation symmetry condition and search the solution as

$$u_n = A_u e^{i(qan-\omega t)} , \quad v_n = A_v e^{i(qan-\omega t)} . \tag{6.45}$$

After substitution to Eqs. (6.44) we get the set of equations for the constants A_i. To formulate these equations it is convenient to express these equations in a matrix form introducing the vector $\mathbf{A} \equiv \begin{pmatrix} A_u & A_v \end{pmatrix}$ and the so-called *dynamic matrix*

$$\hat{D} = \begin{pmatrix} \frac{C_1+C_2}{m_1} & -\frac{C_1+C_2e^{-iaq}}{m_1} \\ -\frac{C_1+C_2e^{iaq}}{m_2} & \frac{C_1+C_2}{m_1} \end{pmatrix} \tag{6.46}$$

The equation for **A** has the form

$$\omega^2 \mathbf{A} - \hat{D}\mathbf{A} = \hat{0}\,. \tag{6.47}$$

This is homogeneous equation; it has a solution only if

$$\det(\omega^2 \hat{1} - \hat{D}) = 0\,. \tag{6.48}$$

This is just the equation which determines the eigenfrequencies. We get

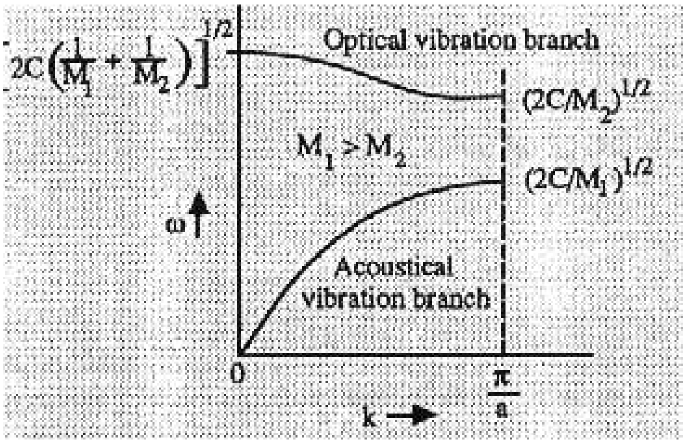

Figure 6.7: Dispersion laws for acoustic and optical phonons at $C_1 = C_2 = C$.

$$\omega_{1,2}^2 = \frac{\omega_0^2}{2}\left[1 \mp \sqrt{1-\gamma^2 \sin^2\frac{aq}{2}}\right] \tag{6.49}$$

where

$$\omega^2 = \frac{(C_1+C_2)(m_1+m_2)}{m_1 m_2}\,, \quad \gamma^2 = 16\left[\frac{C_1C_2}{(C_1+C_2)^2}\right]\left[\frac{m_1m_2}{(m_1+m_2)^2}\right].$$

The frequencies $\omega_{1,2}$ are real because $|\gamma| \le 1$. The spectrum is shown in Fig. 6.7. We see a very important difference with the case of monoatomic chain: there are 2 branches $\omega_{1,2}$ for a given value of q.

The lower branch is called the *acoustic branch* while the upper one is called the *optical branch*. To understand the physical reason for these names let us consider the limits of zero and maximal q. We get

$$\begin{aligned} \omega_{\text{ac}}(0) &= 0\,, & \omega_{\text{ac}}(\pi/a) &= \frac{\omega_0}{\sqrt{2}}\sqrt{1-\sqrt{1-\gamma^2}}\,, \\ \omega_{\text{opt}}(0) &= \omega_0\,, & \omega_{\text{opt}}(\pi/a) &= \frac{\omega_0}{\sqrt{2}}\sqrt{1+\sqrt{1-\gamma^2}}\,. \end{aligned} \tag{6.50}$$

So, we have the inequality chain

$$\omega_{\text{opt}}(0) = \omega_0 > \omega_{\text{opt}}(\pi/a) > \omega_{\text{ac}}(\pi/a) > \omega_{\text{ac}}(0) = 0\,.$$

Now we can discuss the structure of vibrations in both modes. From the dispersion equation (6.47) we get

$$\mathcal{P}_{\text{ac,opt}} = \left(\frac{u_n}{v_n}\right)_{\text{ac,opt}} = \frac{A_u}{A_v} = \frac{C_1 + C_2 e^{-iqa}}{(C_1 + C_2) - m_1 \omega^2_{\text{ac,opt}}} . \tag{6.51}$$

At very long waves ($q \to 0$) we get

$$\mathcal{P}_{ac} = 1 , \quad \mathcal{P}_{opt} = -\frac{m_2}{m_1} \tag{6.52}$$

So, we see that in the acoustic mode all the atoms move next to synchronously, like in an acoustic wave in homogeneous medium. Contrary, in the optical mode; the gravity center remains unperturbed. In an ionic crystal such a vibration produce alternating *dipole moment*. Consequently, the mode is *optical active*. The difference between acoustic and optical modes is shown in Fig. 6.8.

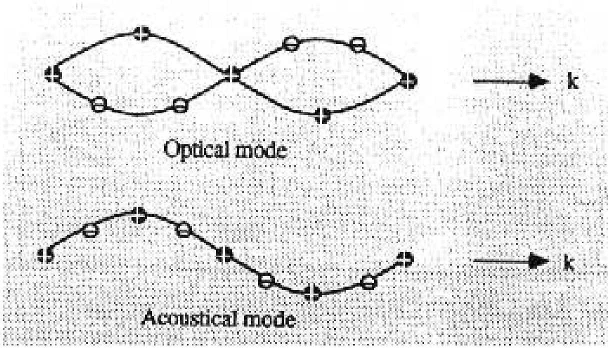

Figure 6.8: Acoustic and optical branches of vibrational spectrum.

Vibration modes of 3D lattices

Now we are prepared to describe the general case of 3D lattice. Assume an elementary cell with s different atoms having masses m_k. Then there are 3 *acoustic* modes with frequencies vanishing at $q \to 0$. At low frequencies they correspond to acoustic waves in an elastic media. One mode is a *longitudinal*, while two others are *transverse* ones. The rest $3s - 3$ ones are the *optical* modes.

Specific Heat of Crystal Lattice

From the classical statistical physics, the average kinetic energy per degree of freedom is

$$\bar{\varepsilon}_k = T/2,$$

the total energy being

$$\bar{\varepsilon} = \bar{\varepsilon}_{\text{pot}} + \bar{\varepsilon}_{\text{kin}} = 2\bar{\varepsilon}_{\text{kin}} .$$

As a result, the total energy is

$$U = 3N_0 \cdot \bar{\varepsilon}_{\text{kin}} = 3RT,$$

where R is the Rydberg constant while N_0 is the Avogadro one. As a result, the specific heat is

$$c_V = 3R = 5.96 \text{ cal/K} \cdot \text{mole}.$$

This relation is violated at low temperatures as far as the temperature becomes less than the so-called *Debye* temperature (which is for most of solids is within the interval 100-400 K), namely it decreases with the temperature decrease. To understand this behavior one should apply quantum mechanics.

First let us recall the average energy for photons calculated previously. The average energy is

$$\bar{\varepsilon} = \frac{\hbar\omega}{2} + \hbar\omega\, N(\omega), \qquad N(\omega) = \frac{1}{e^{\beta\hbar\omega} - 1}$$

where $N(\omega)$ is the *Planck function.* The first item is energy-independent while the second one is just the average energy of bosons with zero chemical potential.

Now we have to recall that the proper excitations in a lattice are collective vibrations, *phonons,* the average energy being

$$U = \sum_{j\mathbf{q}} \hbar\omega_j(\mathbf{q})\, N[\omega_j(\mathbf{q})] = \sum_j \int_0^\infty D_j(\omega)\, \hbar\omega N(\omega)\, d\omega\,.$$

Here j is the number of the vibration branch.

For acoustic modes we use the so-called *Debye model.* Namely, we introduce the average sound velocity s_0 as

$$\frac{1}{s_0^3} = \frac{1}{3}\left(\frac{1}{s_l^3} + \frac{2}{s_t^3}\right)$$

and do the calculation of the density of states similar to the one for photons. As a result, we get

$$VD(\omega) = \frac{3V\hbar\omega^2}{2\pi^3 s_0^3} \equiv V D_0 \omega^2\,.$$

The contribution of acoustic modes to the internal energy is

$$U_{\text{ac}} = V D_0 \int_0^{\omega_D} d\omega\, \frac{\omega^3}{e^{\beta\hbar\omega} - 1}$$

where the so-called Debye frequency is determined from the condition of the total number of modes to be equal to $3N$ for all acoustic branches,

$$3N = V D_0 \int_0^{\omega_D} \omega^2\, d\omega.$$

From this equation

$$\omega_D = s_0 \left(\frac{6\pi^2}{\mathcal{V}_0}\right)^{1/3}, \qquad q_D = \frac{\omega_D}{s_0}$$

where $\mathcal{V}_0$ is the cell volume. The order-of-magnitude estimate for the maximal wave vector q_D is π/a. So according to the so-called *Debye model* all the values of $\mathbf{q}$ are confined in a sphere with the radius q_D. Usually, the *Debye energy* is introduced as

$$\Theta = \hbar\omega_D = \hbar s_0 \left(\frac{6\pi^2}{\mathcal{V}_0}\right)^{1/3}.$$

The typical value of this temperature can be obtained from the rough estimate $a = 10^{-8}$ cm, $s_0 = 10^5$ cm/s. We get $\omega_D = 10^{13}$ s^{-1}, $\Theta = 100$ K. For optical modes, let us assume that $\omega_j(\mathbf{q}) = \omega_{j0}$, i. e. that frequencies of optical phonons are $\mathbf{q}$-independent. Then all the procedure is similar for harmonic oscillators with fixed frequencies. It is conventional also to introduce the energies corresponding to optical branches as $\Theta_j = \hbar\omega_{j0}$. These energies (in degrees) are of the order of 10^2-10^3 K. The quality of the Debye model is demonstrated in Fig. 6.9, where experimental DOS for NaCl is shown together with the Debye approximation.

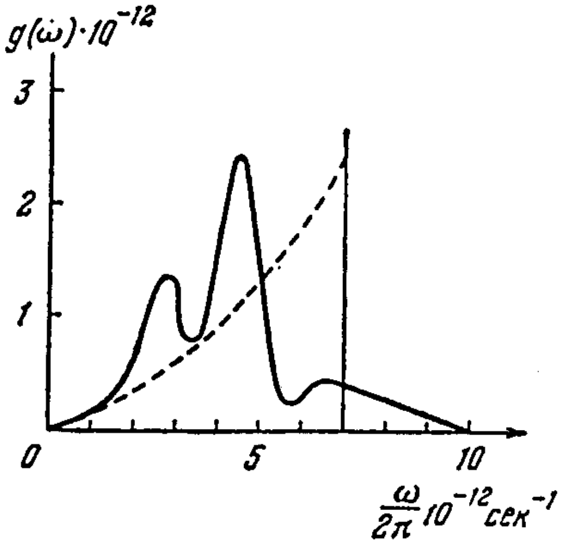

Figure 6.9: On the quality of the Debye approximation for the density of states.

Finally, we get the following expression for the internal energy

$$U = U_0 + NT\left[3\mathcal{D}\left(\frac{\Theta}{T}\right) + \sum_{j=4}^{3s} \frac{\Theta_j/T}{e^{\Theta_j/T} - 1}\right] \tag{6.53}$$

where the *Debye function* $\mathcal{D}(z)$ is

$$\mathcal{D}(z) = \frac{3}{z^3}\int_0^z \frac{x^3\,dx}{e^x - 1}. \tag{6.54}$$

At high temperatures, $T \gg \Theta_j$ (and, consequently, $T \gg \Theta$) we get $z \ll 1$ in Eq. (6.54) and then expand the integrand in powers of x. We see that $\mathcal{D}(0) = 1$. The item under the sum sign in Eq (6.53) are equal to 1, and we get for the sum $3s - 3$. consequently, we get the classical expression

$$\mathcal{E} = \mathcal{E}_0 + 3sNT$$

that leads to the classical expression for the specific heat. At low temperatures we immediately see that optical modes give exponentially small contributions and we can discard them. At the same time, we can replace the upper limit of the integral in (6.54) by infinity. Taking into account that

$$\int_0^\infty \frac{x^3\,dx}{e^x-1} = \frac{\pi^4}{15}$$

we get

$$U = U_0 + \frac{\pi^2 V T^4}{10\hbar^3 s_0^3}$$

that leads to the following expression for the specific heat

$$c_V = \frac{12\pi^4}{5}\left(\frac{T}{\Theta}\right)^3 .$$

The Debye model is very good for low temperatures where only long wave acoustic modes are excited. The acoustic contribution to the specific heat can be expressed through the derivative of the Debye function. We have

$$\frac{c_{ac}}{3N} = 3\left(\frac{T}{\Theta}\right)^3 \int_0^{\Theta/T} \frac{x^4\,dx}{e^x-1} .$$

Temperature dependence of the specific heat according to the Debye model is shown in Fig. 6.10. One can see that really the border between the classical and quantum region corresponds to $T \leq \Theta/3$ rather that to $T \leq \Theta$. The physical reason is strong frequency dependence of phonon DOS. In real life, DOS behavior is much more complicated because of the real band structure effects.

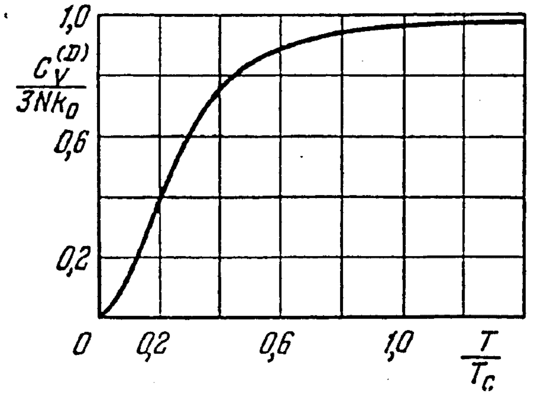

Figure 6.10: Temperature dependence of the specific heat according to the Debye model.

Chapter 7

Statistical ensembles

We have already discussed the Liouville theorem which can be expressed as

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \left[\dot{q}_i \frac{\partial f}{\partial q_i} + \dot{p}_i \frac{\partial f}{\partial p_i} \right] = 0 \, .$$

Because of the Hamilton equations, this equation can be re-written as

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \left[\frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial f}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial f}{\partial p_i} \right] \equiv \frac{\partial f}{\partial t} + \{ f, \mathcal{H} \} = 0 \, .$$

Here the *Poisson brackets* for the quantities $A(p,q)$ and $B(p,q)$ are defined as

$$\{A,B\} \equiv \sum_i \left[\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i} \right] \, . \tag{7.1}$$

Thus, for a stationary distribution we have $\{f, \mathcal{H}\} = 0$. We observe that *any* function of the Hamiltonian $\mathcal{H}$, $f[\mathcal{H}(p,q)]$, satisfies the Liouville equation and can be a proper stationary distribution. Different distributions correspond to different *statistical ensembles*/

7.1 Microcanonical ensemble

As an example, we have already considered the microcanonical ensemble. Let as define $\Gamma(E,N,V)$ as number of states with energies $\leq E$ for given N and V. In the classical mechanics we get

$$\begin{aligned} \Gamma(E,N,V) &\equiv \int_{0<\mathcal{H}<E} \frac{d\Gamma}{(2\pi\hbar)^{3(N_A+N_B+\ldots)} N_A! N_B! \ldots} \\ &= \int \frac{d\Gamma}{(2\pi\hbar)^{3(N_A+N_B+\ldots)} N_A! N_B! \ldots} \Theta[E - \mathcal{H}(p,q)] \end{aligned} \tag{7.2}$$

where

$$\Theta(x) = \begin{cases} 0, & x < 0 \\ 1/2, & x = 0 \\ 1, & x > 0 \end{cases} \tag{7.3}$$

is the Heaviside unit-step function. The many-particle density of states is

$$D(E,N,V)=\frac{\partial\Gamma}{\partial E}=\int\frac{d\Gamma}{(2\pi\hbar)^{3(N_A+N_B+\dots)}N_A!N_B!\dots}\,\delta[E-\mathcal{H}(p,q)]\,. \tag{7.4}$$

Consequently, the normalized distribution function is

$$f(p,q)=\frac{1}{D(E,N,V)}\,\delta[E-\mathcal{H}(p,q)]\,. \tag{7.5}$$

Since the entropy of a normal system is given by Eq. (3.31), $S(E,N,V)\approx\log\Gamma(E,N,V)$, we obtain

$$\frac{1}{T}=\frac{\partial}{\partial E}\log\Gamma=\frac{1}{\Gamma}\frac{\partial\Gamma}{\partial E}=\frac{D(E,N,V)}{\Gamma(E,N,V)}\,.$$

Now we can re-derive the equipartition principle for ideal gas as follows. Consider the auxiliary relationship,

$$\begin{aligned}\left\langle p_n\frac{\partial\mathcal{H}}{\partial p_n}\right\rangle &= \frac{1}{D}\int\frac{d\Gamma}{(2\pi\hbar)^{3N}N!}\,p_n\frac{\partial\mathcal{H}}{\partial p_n}\,\delta(E-\mathcal{H})\\ &= -\frac{1}{D}\int\frac{d\Gamma}{(2\pi\hbar)^{3N}N!}\,p_n\frac{\partial}{\partial p_n}\,\Theta(E-\mathcal{H}) \quad\rightarrow\quad \text{integration by parts}\rightarrow\\ &= \frac{1}{D}\int\frac{d\Gamma}{(2\pi\hbar)^{3N}N!}\,\Theta(E-\mathcal{H})=\frac{\Gamma}{D}=T\,.\end{aligned}$$

Using this relation, we obtain

$$K=\left\langle\frac{p_n^2}{2m}\right\rangle=\frac{T}{2}\,.$$

Another quantity which we shall need later is the so-called *virial*,

$$\theta=-\sum_n q_n\dot{p}_n=\sum_n q_n\frac{\partial\mathcal{H}}{\partial q_n}\,.$$

For a 3D system with *harmonic* potential energy, $V\propto\sum_{ik}\alpha_{ik}q_iq_k$, one can easily prove that

$$\langle\theta\rangle=3NT\,.$$

Free particles: Let us demonstrate the general principles using the gas of free particles. We have to be able to calculate the volume of the hyper-sphere,

$$V_n\equiv\int\prod_{i=1}^{n}dp_i\Theta\left(E-\sum_{i=1}^{3N}\frac{p_i^2}{2m}\right)=(2mE)^{3N/2}\mathcal{V}_n\,, \tag{7.6}$$

where

$$\mathcal{V}_n=\int_{-\infty}^{\infty}\prod_i dx_i\,\Theta\left(1-\sum_i x_i^2\right)\,.$$

To calculate $\mathcal{V}_n$ let us introduce an auxiliary function

$$\tilde{\mathcal{V}}_n(R) = \int_{-\infty}^{\infty} \prod_i dx_i \, \Theta\left(R^2 - \sum_i x_i^2\right), \quad \mathcal{V}_n = \tilde{\mathcal{V}}_n(1).$$

consider an auxiliary integral

$$I_n \equiv \int_{-\infty}^{\infty} \prod_{i=1}^{n} dx_i \, e^{-x_i^2} = \pi^{n/2}. \tag{7.7}$$

This integral can be also written as

$$I_n = \int d\tilde{\mathcal{V}}_n e^{-R^2}, \quad R^2 = \sum_i x_i^2. \tag{7.8}$$

Since $\tilde{\mathcal{V}}_n = \mathcal{V}_n R^n$ we have $d\tilde{\mathcal{V}}_n = n\mathcal{V}_n R^{n-1}$. Comparing Eqs. (7.7) and (7.8) we get

$$\pi^{n/2} = n\mathcal{V}_n \int_0^{\infty} dR\, R^{n-1} e^{-R^2} \quad \to \quad \mathcal{V}_n = \frac{\pi^{n/2}}{(n/2)!}.$$

Consequently, we recover old results,

$$\begin{aligned} \Gamma(E) &= \frac{V^N}{(2\pi\hbar)^{3N}} \frac{\pi^{3N/2}}{(3N/2)!} (2mE)^{3N/2}; \\ D(E) &= \frac{\partial \Gamma}{\partial E} = \frac{V^N}{(2\pi\hbar)^{3N}} \frac{\pi^{3N/2}}{(3N/2-1)!} (2mE)^{3N/2-1}; \\ S(E) &= N\left\{\frac{5}{2} + \log\left[\frac{V}{N}\left(\frac{mE}{3\pi N\hbar^2}\right)^{3/2}\right]\right\}, \end{aligned}$$

etc. Then, we recover the equation of state.

7.2 Canonical ensembles

Suppose that there are other than the energy E integrals of motion,

$$C_i(q,p) = C_i = \text{const}.$$

Since $\{C, \mathcal{H}\} = 0$ any function $f[\mathcal{H}(q,p), C_i(q,p)]$ can be a proper stationary distribution function. Different choice of the system gives rise to different ensembles.

Canonical ensemble

In this ensemble, the temperature, volume, and number of particles are fixed. Moving along the lines of Gibbs' method we have obtained the canonical distribution (Sec. 5.2)

$$f_l = \frac{1}{Z} e^{-\beta E_l}\,, \quad Z = \sum_s e^{-\beta E_s}\,.$$

In classical mechanics,

$$f(p,q) = \frac{1}{Z} e^{-\beta \mathcal{H}(p,q)}\,,$$

where Z is defined by the normalization condition. Here the temperature is fixed, while the energy fluctuate around the average energy U. The fluctuations are given by the relation

$$\begin{aligned} (\Delta E)^2 &\equiv \langle E^2\rangle - \langle E\rangle^2 = \frac{1}{Z}\sum_s E_s^2 e^{-\beta E_s} - \left[\frac{1}{Z}\sum_s E_s e^{-\beta E_s}\right]^2 \\ &= \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} - \left(\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right)^2 = \frac{\partial}{\partial \beta}\left(\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right) = -\frac{\partial U}{\partial \beta} = T^2 C_V\,. \end{aligned}$$

Grand canonical ensemble

Now let as allow an exchange by particles with the surrounding medium. Along the Gibbs method, we have obtained (see Sec. 5.2.1)

$$f_{Nl} = \frac{1}{\Xi} e^{\beta(\mu N - E_{Nl})}\,, \quad \Xi = \sum_{N=0}^{\infty} e^{\beta\mu N} \sum_l e^{-\beta E_{Nl}} = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_N\,.$$

Again,

$$\begin{aligned} U &= \frac{1}{\Xi}\sum_{Ns} E_{Ns} e^{\beta(\mu N - E_{Ns})} = -\frac{\partial \log \Xi}{\partial \beta}\,; \\ S &= -\sum_{Ns} f_{Ns} \log f_{Ns} = -\beta(\mu \bar{N} - U) - \log \Xi\,. \end{aligned}$$

Here we have introduced the average particle number,

$$\bar{N} \equiv \sum_{Ns} N f_{Ns} = \frac{1}{\Xi}\sum_{N=0}^{\infty} N e^{\beta\mu N} Z_N = \frac{T}{\Xi}\frac{\partial \Xi}{\partial \mu}\,. \tag{7.9}$$

Introducing

$$\Omega \equiv -T \log \Xi$$

we get

$$\bar{N} = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V}\,.$$

Then is becomes obvious that Ω coincides with the Landau free energy since

$$\Omega = U - TS - \mu\bar{N} = -PV\,.$$

It is easy to derive the fluctuation in the particle number (check!),

$$(\Delta N)^2 = \langle N^2\rangle - \bar{N}^2 = -T\left(\frac{\partial^2\Omega}{\partial\mu^2}\right)_{T,V}\,.$$

The latter expression can be converted into a more convenient form introducing the gas density, $\rho = N/V$, using the expression for $\bar{N}$,

$$\rho \equiv \frac{\bar{N}}{V} = \left(\frac{\partial P}{\partial\mu}\right)_T\,.$$

Then we obtain,

$$(\Delta N)^2 = TV\left(\frac{\partial\rho}{\partial\mu}\right)_T = TV\left(\frac{\partial\rho}{\partial P}\right)_T\left(\frac{\partial P}{\partial\mu}\right)_T = TN\rho K_T\,.$$

Poisson distribution. An instructive exercise is to find the distribution of the particle number in a given volume V of an ideal gas. Since $f_N \propto e^{\beta\mu N}\sum_l e^{-\beta E_{Nl}} \propto e^{\beta\mu N}Z_N$ we get

$$f_N = \frac{Z_N}{\Xi}e^{\beta\mu N}, \quad \Xi = \sum_{N=0}^{\infty} e^{\beta\mu N}Z_N\,.$$

For an ideal gas,

$$Z_N = \frac{Z_1^N}{N!} = \frac{1}{N!}\left(\frac{V}{\Lambda^3}\right)^N, \quad \Xi = \sum_{N=0}^{\infty} e^{\beta\mu N}Z_N = \exp\left(\frac{V}{\Lambda^3}e^{\beta\mu}\right)\,.$$

Now, from Eq. (7.9) we have

$$\bar{N} = \frac{\partial\log\Xi}{\partial\beta\mu} = (V/\Lambda^3)e^{\beta\mu} \quad \to \quad \Xi = e^{\bar{N}},\ Z_1^N e^{\beta\mu N} = \bar{N}^N\,.$$

Finally,

$$f_N = \frac{\bar{N}^N}{N!}e^{-\bar{N}}\,. \tag{7.10}$$

that is the famous Poisson distribution.

Exercise. Find $\overline{\Delta N} \equiv \sqrt{\langle\Delta N^2\rangle}$ from the Poisson distribution. Derive an approximate expression for Eq. (7.10) at $\bar{N} \gg 1,\ |N-\bar{N}| \ll \bar{N}$.

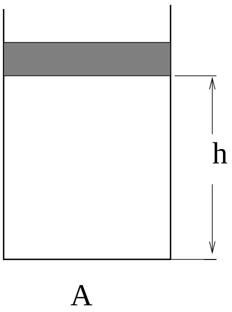

Figure 7.1: On canonical ensemble with a fixed pressure.

Other ensembles

Consider the canonical ensemble with a fixed pressure rather that volume, the so-called *pressure ensemble*. The system is shown in Fig. 7.1. Here the integral of motion is the *enthalpy* rather that the energy. Then we come to the distribution function

$$f(q,p)=\frac{1}{Z_P}e^{-\beta\mathcal{H}_P}\,,\quad \mathcal{H}_P=\mathcal{H}(p,q)+\frac{k^2}{2M}+PAh\,.$$

Here k is the piston momentum while M is its mass. Z_P is again the normalization factor. Similarly, we can define

$$H=U+P\bar{V}=-\frac{\partial}{\partial\beta}\log Z_P\,,\quad \bar{V}=-\frac{\partial}{\partial\beta}\log Z_P\,.$$

One can easily check that

$$G=-T\log Z_P\,.$$

Now we can integrate out the kinetic energy of a piston an replace $dh\to dV/A$ to get

$$dh\int\frac{dk}{2\pi\hbar}e^{-\beta k^2/2M}=\frac{dV}{V_M}\,,\quad V_M=A\Lambda_M=A\sqrt{\frac{2\pi\hbar^2}{MT}}\,.$$

As a result,

$$Z_P=\int\frac{dV}{V_M}e^{-\beta PV}Z_N(V,T)\,,$$

where $Z_N(V,T)$ is the partition function calculated for fixed volume and fixed number of particles. We observe that the pressure ensemble can be considered as Laplace-transformed canonical ensemble for fixed volume. Since for an ideal gas $Z_N\propto V^N$ the distribution function for the volume occupied by the gas under a given pressure is

$$f_V=CV^Ne^{-\beta PV}\,,$$

where C is the normalization constant (find it and express through $\bar{V}$!)

We can apply the above considerations to microcanonical and canonical ensembles to obtain the relation

$$Z_N(V,T)=\int dE\,e^{-\beta E}D(E)\,,$$

where $D(E)$ is the partition function for the microcanonical ensemble.

To find the energy distribution in the canonical ensemble let us expand the function

$$u(E) \equiv -\log[D(E)e^{-\beta E}] = \beta E - S(E)$$

up to the second order in the difference $E-U$ and using the relation

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial \beta}{\partial E} = -\beta^2/2C_V$$

we obtain the energy probability distribution as

$$f_E \propto \exp\left[-\frac{\beta^2(E-U)^2}{2C_V}\right], \quad (\Delta E)^2 = T^2 C_V\,. \tag{7.11}$$

Exercise. Calculate the normalization constant.

7.3 Ensembles in quantum statistics

Pure and mixed ensembles

Let us consider a particle with spin $S=1/2$. Quantizing spin along z axis, we find 2 basic states

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence,

$$\hat{S}_z|\uparrow\rangle = \frac{1}{2}|\uparrow\rangle, \quad \hat{S}_z|\downarrow\rangle = -\frac{1}{2}|\downarrow\rangle\,.$$

Consequently, the operator S_z in this representation can be expressed as

$$\hat{S}_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

while the conjugated wave functions are just the strings

$$\langle\uparrow| = (1\ 0), \quad \langle\downarrow| = (0\ 1)\,.$$

Consider now the the ensemble where the particles are quantized along different direction. Still the state can be expanded in eigenfunctions as

$$|\theta\rangle = \cos\frac{\theta}{2}|\uparrow\rangle + \sin\frac{\theta}{2}|\downarrow\rangle\,.$$

Then we easily obtain,

$$\langle\hat{S}_z\rangle = \frac{1}{2}\left(\cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}\right) = \frac{1}{2}\cos\theta\,.$$

This state corresponds to the system polarized along the axis

$$\mathbf{n} = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y\,.$$

Here $\mathbf{e}$ is the unit vector. The operator $\hat{S}_x$ in the representation used has the form

$$\hat{S}_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We can easily show that

$$\langle \hat{S}_x \rangle = \frac{1}{2}\sin\theta\,.$$

The above example where all the spins are polarized along a given axis is called a *pure* state. If the spins occupy different states with a given probability, the ensemble is called *mixed*. To see the difference with a pure ensemble let us consider a system there the probabilities for the states $|\uparrow\rangle$ and $|\downarrow\rangle$ are equal. Then,

$$\langle \hat{S}_\alpha \rangle = \sum_{i=\uparrow,\downarrow} w_i \langle i|\hat{S}_\alpha|i\rangle = 0\,,$$

for any direction, $\langle \hat{\mathbf{S}}\cdot\mathbf{n}\rangle = 0$. For different probabilities w_i this average is finite, but $\leq 1/2$. The maximum is reached only for a pure state where one of the probabilities is zero.

Now we can describe the general picture for a mixed ensemble. The ensemble average is now

$$\langle \hat{A} \rangle = \sum_i w_i \langle i|\hat{A}|i\rangle\,.$$

Expanding the states $|i\rangle$ is terms of any orthonormal basis,

$$|i\rangle = \sum_\alpha c_i |\alpha\rangle$$

we rewrite the average as

$$\langle \hat{A} \rangle = \sum_i w_i \sum_{\alpha\beta} c^*_{\beta i} c_{\alpha i}\,\langle\beta|\hat{A}|\alpha\rangle \equiv \sum_{\alpha\beta} A_{\beta\alpha}\,\rho_{\alpha\beta}\,.$$

Here $A_{\beta\alpha} \equiv \langle\beta|\hat{A}|\alpha\rangle$is the matrix element of the operator $\hat{A}$, while

$$\rho_{\alpha\beta} = \sum_i w_i c^*_{\beta i} c_{\alpha i} = \sum_i w_i \langle\alpha|i\rangle\langle i|\beta\rangle$$

is called the density matrix. It can be expressed as a matrix element of the *density operator*,

$$\hat{\rho} = \sum_i w_i\,|i\rangle\langle i|\,.$$

Now the average can be expressed as

$$\langle \hat{A} \rangle = \mathrm{Tr}\,\hat{A}\,\hat{\rho}\,. \tag{7.12}$$

For a pure state,

$$\hat{\rho} = |\psi\rangle\langle\psi|\,, \quad \hat{\rho}^2 = \hat{\rho}\,,$$

thus the density operator can have the eigenvalues 0 or 1.

Density operator for spin 1/2

Any 2×2 matrix can be generally expanded in the unit matrix, $\hat{1}$ and the *Pauli matrices*

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In general, we can write,

$$\hat{\rho} = \frac{1}{2}\left(\hat{1} + \hat{\sigma} \cdot \mathbf{P}\right),$$

where $\hat{1}$ is the unit matrix, $\hat{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, while $\mathbf{P}$ is called the *polarization vector*. We can easily show that $\mathrm{Tr}\ \hat{\rho} = 1$ (since $\mathrm{Tr}\ \hat{1} = 2$ and $\mathrm{Tr}\ \hat{\sigma}_i = 0$). For a pure state one has to require

$$\mathbf{P} = \cos\theta\, \mathbf{e}_x + \sin\theta\, \mathbf{e}_y\,,$$

while for a mixed ensemble $P = |\mathbf{P}| < 1$. P is called the degree of polarization.

Density matrix for thermodynamic ensembles

Since

$$i\hbar \frac{\partial}{\partial t}|i\rangle = \hat{\mathcal{H}}|i\rangle$$

we can show that

$$\frac{d\hat{\rho}}{dt} = \frac{\partial \hat{\rho}}{\partial t} + \frac{i}{\hbar}\left[\hat{\rho}, \hat{\mathcal{H}}\right] = 0\,.$$

Here $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the commutator which in classical physics corresponds to the Poisson brackets. Now we can construct statistics similarly to the classical one replacing distribution functions by density operators. In this way, for canonical ensemble we obtain

$$\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{\mathcal{H}}}, \quad Z = \mathrm{Tr}\, e^{-\beta \hat{\mathcal{H}}} = \sum_{\alpha} e^{-\beta E_\alpha}\,.$$

The grand canonical ensemble is defined in a similar way,

$$\hat{\rho} = \frac{1}{\Xi} e^{-\beta(\hat{\mathcal{H}} - \mu \hat{N})}\,.$$

For a microcanonical ensemble, in the representation of eigenstates $|\alpha_i\rangle$ belonging to ith level the Hamiltonian can be written as

$$\hat{\mathcal{H}} = \sum_i E_i \hat{\Pi}_i\,, \quad \hat{\Pi}_i = \sum_{\alpha_i} |\alpha_i\rangle\langle\alpha_i|\,.$$

Since $\hat{\Pi}_i \hat{\Pi}_j = \delta_{ij} \hat{\Pi}_i$ the projection operator Π_i commutes with the Hamiltonian. Thus, for the system having the energy E_i

$$\hat{\rho} = \frac{1}{W_i} \hat{\Pi}_i\,, \quad W_i = \mathrm{Tr}\, \hat{\Pi}_i\,.$$

As usual,

$$S = -\mathrm{Tr}\, \hat{\rho}\, \log \hat{\rho}\,.$$

Examples

Spin 1/2 in a magnetic field:

$$\hat{\mathcal{H}} = -\mu \cdot \mathbf{B} = -\mu B \sigma_z \,,$$

where μ is the magnetic moment. Since

$$e^{-\beta\hat{\mathcal{H}}} = \sum_n \frac{(-\beta\hat{\mathcal{H}})^n}{n!} = \sum_n \frac{(\beta\mu B\sigma_z)^n}{n!} = \cosh(\beta\mu B) + \sigma_z \sinh(\beta\mu B)$$

(here we employ the fact that $\sigma_z^{2k} = 1$, $\sigma_z^{2k+1} = \sigma_z$) we obtain

$$\hat{\rho} = \frac{1}{Z} e^{-\beta\hat{\mathcal{H}}} = \frac{1}{2}\left[\hat{1} + \hat{\sigma}_z \tanh(\beta\mu B)\right] \,, \quad \langle\sigma_z\rangle = \operatorname{Tr}\hat{\rho}\hat{\sigma}_z = \tanh\beta\mu B.$$

A free particle in a box with volume V:

$$\hat{\mathcal{H}} = \hat{\mathbf{p}}^2/2m \quad \to \quad \langle\mathbf{r}|\hat{\rho}|\mathbf{r}'\rangle = \frac{1}{Z}\sum_{\mathbf{p},\mathbf{p}'} \langle\mathbf{r}|\mathbf{p}'\rangle\langle\mathbf{p}'|e^{-\beta\hat{\mathbf{p}}^2/2m}|\mathbf{p}\rangle\langle\mathbf{p}|\mathbf{r}'\rangle \,.$$

Since

$$\langle\mathbf{r}|\mathbf{p}\rangle = \frac{1}{\sqrt{V}} e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}} \,, \quad \langle\mathbf{p}'|e^{-\beta\hat{\mathbf{p}}^2/2m}|\mathbf{p}\rangle = \delta_{\mathbf{p}\mathbf{p}'} e^{-\beta p^2/2m}$$

we get

$$\langle\mathbf{r}|\hat{\rho}|\mathbf{r}'\rangle = \frac{1}{Z}\int \frac{d^3p}{(2\pi\hbar)^3} e^{-\beta p^2/2m + (i/\hbar)\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')} = \frac{1}{Z}\left(\frac{m}{2\pi\beta\hbar^2}\right)^{3/2} e^{-(m/2\beta\hbar^2)|\mathbf{r}-\mathbf{r}'|^2} \,.$$

From the normalization conditions we obtain

$$Z = \operatorname{Tr} e^{-\beta\hat{\mathcal{H}}} = \int d^3r \, \langle\mathbf{r}|e^{-\beta\hat{\mathcal{H}}}|\mathbf{r}\rangle = \frac{V}{\Lambda^3} \,, \quad \langle\mathbf{r}|\hat{\rho}|\mathbf{r}'\rangle = \frac{1}{V} e^{-\pi|\mathbf{r}-\mathbf{r}'|^2/\Lambda^2} \,.$$

Chapter 8

Fluctuations

In this section we discuss some additional aspect regarding fluctuations.

8.1 The Gaussian distribution

Let us consider a closed system described by a microcanonical ensemble.

Assume that x is some quantity of the subsystem, $\bar{x} = 0$. At this point entropy has a maximum,

$$\frac{\partial S}{\partial x} = 0\,, \quad \frac{\partial^2 S}{\partial x^2} < 0\,.$$

Thus one can expand

$$S(x) = S(\bar{x}) - \gamma \frac{x^2}{2}\,,$$

the normalized distribution being

$$w(x)\,dx = \sqrt{\frac{\gamma}{2\pi}} e^{-\gamma x^2/2}\,dx\,. \tag{8.1}$$

As

$$\langle x^2 \rangle = \gamma^{-1}\,, \tag{8.2}$$

we can write the result (the so-called Gaussian distribution) in the final form

$$w(x)\,dx = \frac{1}{\sqrt{2\pi \langle x^2 \rangle}} \exp\left(-\frac{x^2}{2\langle x^2 \rangle} \right) dx\,. \tag{8.3}$$

As x is small, for any quantity ϕ

$$\langle (\Delta\phi)^2 \rangle = \left[\frac{d^2\phi}{dx^2} \right]_{x=0} \langle x^2 \rangle\,.$$

For several (n) variables we have

$$w = Ae^{-\sum_{ik}\gamma_{ik}x_i x_k/2}\,, \qquad \int w \prod_i dx_i = 1\,.$$

The last equality determines the normalization constant A. In this way we get

$$w(x) = \frac{\sqrt{\gamma}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{ik}\gamma_{ik}x_i x_k\right)\,, \qquad \gamma \equiv \det\{\gamma_{ik}\}\,.$$

Now let us define *thermal conjugates*

$$X_i = -\frac{\partial S}{\partial x_i} = \sum_k \gamma_{ik} x_k\,.$$

The averages $\langle x_i X_k \rangle$ can be easily calculated:

$$\langle x_i X_k \rangle = \frac{\sqrt{\gamma}}{(2\pi)^{n/2}} \sum_l \int x_i \gamma_{kl} x_l \exp\left(-\frac{1}{2}\sum_{pq}\gamma_{pq}x_p x_q\right) \prod_{s=1}^{n} dx_s\,.$$

One can prove that it is equal to δ_{ik},[1]

$$\langle x_i X_k \rangle = \delta_{ik}\,.$$

From that we get

$$\langle x_i x_k \rangle = (\hat{\gamma}^{-1})_{ik}\,, \quad \langle X_i X_k \rangle = \gamma_{ik}\,.$$

Thus

$$\langle (\Delta\phi)^2 \rangle = \frac{\partial\phi}{\partial x_i}\frac{\partial\phi}{\partial x_k}(\hat{\gamma}^{-1})_{ik}\,.$$

For *statistically-independent* fluctuations

$$(\hat{\gamma}^{-1})_{ik} = 0\,.$$

Thus the distributions are uncoupled.

---

[1] **Proof:** Assume for a moment that the average $\bar{x_i} = x_{i0} \neq 0$. Then

$$x_{i0} = \frac{\sqrt{\gamma}}{(2\pi)^{n/2}} \int \prod_i dx_i\, x_i\, e^{-(1/2)\sum_{st}\gamma_{st}(x-x_{s0})(x-x_{t0})}\,.$$

After differentiation of this equation with respect to x_{x0} and putting $x_{i0} = 0$, $x_{k0} = 0$ we obtain the above result.

8.2 Fluctuations of fundamental thermodynamic quantities

As the change in the entropy is proportional to *minimal work* to change the thermodynamic properties *reversible* the probability for a given configuration can be written as, see Eq. (4.29),

$$w \propto e^{-W_{\min}/T} \propto \exp\left(-\frac{\Delta U - T\,\Delta S + P\,\Delta V}{T}\right).$$

Regarding that all the quantities are the functions of S and V and expanding U in small deviations ΔS and ΔV from the equilibrium values,

$$\Delta U = T\Delta S - P\Delta V + \frac{1}{2}\left[\frac{\partial^2 U}{\partial S^2}(\Delta S)^2 + 2\frac{\partial^2 U}{\partial S \partial V}\Delta S\,\Delta V + \frac{\partial^2 U}{\partial V^2}(\Delta V)^2\right]$$

we get

$$W_{\min} = \frac{1}{2}\left[\Delta S\,\Delta\left(\frac{\partial U}{\partial S}\right)_V + \Delta V\,\Delta\left(\frac{\partial U}{\partial V}\right)_S\right] = \frac{1}{2}(\Delta S\,\Delta T - \Delta P\,\Delta V)$$

Thus

$$w \propto \exp\left(\frac{\Delta P\,\Delta V - \Delta S\,\Delta T}{2T}\right). \tag{8.4}$$

This is the general formula to find fluctuations of various thermodynamic quantities, and below we give several examples in which different parameters are considered as independent.

Examples

Independent V and T:

$$\begin{aligned}
\Delta S &= \left(\frac{\partial S}{\partial T}\right)_V \Delta T + \left(\frac{\partial S}{\partial V}\right)_T \Delta V = \frac{C_V}{T}\Delta T + \left(\frac{\partial P}{\partial T}\right)_V \Delta V\,,\\
\Delta P &= \left(\frac{\partial P}{\partial T}\right)_V \Delta T + \left(\frac{\partial P}{\partial V}\right)_T \Delta V\,.
\end{aligned}$$

Substituting these expressions in Eq. (8.4), we get

$$w \propto \exp\left\{-\frac{C_V}{2T^2}(\Delta T)^2 + \frac{1}{2T}\left(\frac{\partial P}{\partial V}\right)_T (\Delta V)^2\right\}.$$

Thus we reproduce previous results:

$$\langle \Delta T\,\Delta V\rangle = 0\,, \quad \langle(\Delta T)^2\rangle = T^2/C_V\,, \quad \langle(\Delta V)^2\rangle = -T(\partial V/\partial P)_T\,. \tag{8.5}$$

Independent P and S:

$$\begin{aligned}
\Delta V &= \left(\frac{\partial V}{\partial P}\right)_S \Delta P + \left(\frac{\partial V}{\partial S}\right)_P \Delta S\,, \\
\Delta T &= \left(\frac{\partial T}{\partial S}\right)_P \Delta S + \left(\frac{\partial T}{\partial P}\right)_S \Delta P = \frac{T}{C_P}\Delta S + \left(\frac{\partial T}{\partial P}\right)_S \Delta P\,.
\end{aligned}$$

As

$$dH = T\,dS + V\,dP$$

we have

$$\left(\frac{\partial V}{\partial S}\right)_P = \frac{\partial^2 H}{\partial P \partial S} = \left(\frac{\partial T}{\partial P}\right)_S$$

and

$$\Delta V = \left(\frac{\partial V}{\partial P}\right)_S \Delta P + \left(\frac{\partial T}{\partial P}\right)_S \Delta S\,.$$

Thus

$$w \propto \exp\left\{\frac{1}{2T}\left(\frac{\partial V}{\partial P}\right)_S (\Delta P)^2 - \frac{1}{2C_P}(\Delta S)^2\right\}\,.$$

As a result,

$$\langle \Delta S \Delta P\rangle = 0\,, \quad \langle (\Delta S)^2\rangle = C_P\,, \quad \langle (\Delta P)^2\rangle = -T\left(\frac{\partial P}{\partial V}\right)_S\,.$$

Fluctuations per particle: From Eq. (8.5) we get:

$$\left\langle \Delta\left(\frac{V}{N}\right)^2\right\rangle = -\frac{T}{N^2}(\partial V/\partial P)_T\,.$$

Putting $V = \text{const}$ we obtain

$$\langle (\Delta N)^2\rangle = -\frac{TN^2}{V^2}(\partial V/\partial P)_T$$

This formula can be re-written as

$$-\frac{N^2}{V^2}\left(\frac{\partial V}{\partial P}\right)_{T,N} = N\left(\frac{\partial}{\partial P}\frac{N}{V}\right)_{T,N}\,.$$

Then

$$N\left(\frac{\partial}{\partial P}\frac{N}{V}\right)_{T,N} = \frac{N}{V}\left(\frac{\partial N}{\partial P}\right)_{T,V} = \left(\frac{\partial N}{\partial P}\right)_{T,V}\left(\frac{\partial P}{\partial \mu}\right)_{T,V} = \left(\frac{\partial N}{\partial \mu}\right)_{T,V}\,.$$

Here we have employed the fact that the ratio N/V depends only on P and T (additivity principle). The second equality,

$$\frac{N}{V} = \left(\frac{\partial P}{\partial \mu}\right)_{T,V}$$

follows from the formula

$$d\Omega = -V\,dP = -S\,dT - N\,d\mu .$$

Finally we arrive at a very useful formula

$$\langle (\Delta N)^2 \rangle = T\left(\frac{\partial N}{\partial \mu}\right)_{T,V} .$$

8.3 Correlation of fluctuations in time. Langevin method

Consider correlation between the values of the quantity x ($\bar{x} = 0$) at different instants. Such correlation is characterized by the average $\langle x(t)x(t')\rangle$. Since under stationary external conditions the average depends only on the difference $t'-t$ one can define

$$\phi(t) \equiv \langle x(0)x(t)\rangle .$$

It is clear from obvious symmetry considerations that

$$\phi(t) = \phi(-t) .$$

A proper definition for quantum mechanical case is

$$\phi(t-t') = \frac{1}{2}\langle \hat{x}(t)\hat{x}(t') + \hat{x}(t')\hat{x}(t)\rangle .$$

Now let us consider a state of partial equilibrium characterized by a given value x. For simplicity let us also assume that x is a *classical quantity*.

Now we are prepared to discuss the state far from the equilibrium, namely assume that $x \gg \sqrt{\langle x^2\rangle}$. If we assume that the state under consideration is determined *only* by the quantity x we can suggest that the rate $\dot{x}$ for restoring the equilibrium is some function of one variable x. The simplest function valid for small enough x is the linear one,

$$dx/dt = -\lambda x , \quad \lambda > 0 . \tag{8.6}$$

It is natural to assume that this equation is the same as the “equation of motion” for a macroscopic quantity $\bar{x}$. To express this concept in a more mathematical form let us define the quantity $\xi_x(t)$ as the mean value of x at time $t > 0$ under condition that at $t = 0$ its value was x. The quantity $\xi_x(t)$ is defined through the *conditional probability* $P(x',t'|x,t)$ to find the the value x' of the variable x under condition that at time t it was x. We have,

$$\xi_x(t) = \int dx'\, x' P(x',t'|x,t) . \tag{8.7}$$

Consequently,

$$\phi(t) = \langle x\xi_x(t)\rangle = \int dx\, w(x)\, x\, \xi_x(t) , \quad t > 0 .$$

For the values $\xi_x \gg \sqrt{\langle x^2 \rangle}$ we obtain

$$d\xi_x(t)/dt = -\lambda \xi_x(t)\,, \quad t > 0\,.$$

Consequently,

$$\xi_x(t) = x e^{-\lambda t} \quad \to \quad \phi(t) = \langle x^2 \rangle e^{-\lambda t} = \gamma^{-1} e^{-\lambda t}\,, \quad t > 0\,.$$

Here γ is the constant which enters the Gaussian distribution, see Eq. (8.2). In general, at

$$\phi(t) = \gamma^{-1} e^{-\lambda |t|}\,. \tag{8.8}$$

It the random quantity x is comparable with the mean square fluctuation one has to add the *random force y* to the “equation of motion” (8.6) for the fluctuation x,

$$\dot{x} = -\lambda x + y(t)\,.$$

The correlation function $\langle y(0) y(t) \rangle$ must be defined in a way to provide correct expression (8.8) for $\langle x(0) x(t) \rangle$ at large x. To do that we find a formal solution of the above equation with some initial condition,

$$x(t) = e^{-\lambda t} \int_{t_0}^{t} d\tau\, e^{\lambda \tau} y(\tau)\,.$$

The average $\langle x(t) x(t') \rangle$ is then

$$\langle x(t) x(t') \rangle = e^{-\lambda(t+t')} \int_{t_0}^{t} dt_1 \int_{t_0}^{t'} dt_2\, e^{\lambda(t_1+t_2)} \langle y(t_1) y(t_2) \rangle\,.$$

This result must be independent of t_0 provided $\lambda(t - t_0) \gg 1$ and $\lambda(t' - t_0) \gg 1$. Thus we can put $t_0 = -\infty$. The only way to obtain $\langle x(t) x(t') \rangle$ dependent only on $|t - t'|$ is to suggest that

$$\langle y(t_1) y(t_2) \rangle = C \delta(t_1 - t_2)\,.$$

Substituting this expression into Eq. (8.8) for $\langle x(t) x(t') \rangle$ for large t we get

$$\langle y(t_1) y(t_2) \rangle = (2\lambda/\gamma)\, \delta(t_1 - t_2)\,.$$

The above result can be easily generalized for the case of several fluctuating quantities, x_k. Similarly, we define

$$\phi_{ik}(t' - t) = \langle x_i(t') x_k(t) \rangle\,, \quad \phi_{ik}(t) = \phi_{ki}(-t)\,.$$

However, there is one more symmetry property which is a consequence of the invariance with respect to time reversal.[2] Because of the time-reversal symmetry,

$$\phi_{ik}(t) = \phi_{ik}(-t) = \phi_{ki}(t)\,.$$

---

[2] The system is assumed to be not in a magnetic field and not rotating as a whole.

There is an important remark here. We have tacitly assumed that the quantities x_k are not affected directly by time reversal. However, these quantities can change sign under time reversal, like the quantities proportional to velocities. If both x_i and x_k possess such a property the above equation is valid, of only one of two changes its sign, then there is an additional "-" sign.

The following generalization is straightforward. We introduce the matrix λ_{ik} instead of the scalar λ. The fluctuations x_k are to be determined from the set of differential equations

$$\dot{x}_i = -\sum_k \lambda_{ik} x_k + y_i(t) \tag{8.9}$$

with random forces y_i defined by the correlation properties

$$\langle y_i(t_1) y_k(t_2) \rangle = 2\eta_{ik} \delta(t_1 - t_2)\,, \quad \hat{\eta} = \hat{\lambda}\hat{\gamma}^{-1}\,. \tag{8.10}$$

Matrix elements of $\hat{\eta}$ are called the *kinetic coefficients*. Their physical meaning will be discussed in the next section. The scheme discussed above is called the *Langevin method*. Again,

$$\phi_{ik}(0) = (\hat{\gamma}^{-1})_{ik}\,. \tag{8.11}$$

Simple example: Brownian motion

Consider a particle moving in a liquid. It is subjected to collisions with liquid molecules. Due to such collisions the particle's velocity $\mathbf{v}$ changes in time according to the equation

$$m\dot{v}_i = -\nu v_i + f_i(t)$$

where ν is the viscous friction coefficient while f_i are random forces. Consequently,

$$\hat{\lambda} = (\nu/m)\,\delta_{ik}\,, \quad y_i = f_i/m\,, \quad \hat{\gamma}^{-1} = \langle v_i^2(0)\rangle \delta_{ik} = \frac{1}{3}\langle v^2\rangle\,\delta_{ik}\,.$$

Let us now consider a displacement $\mathbf{x}(t) = \mathbf{r}(t) - \mathbf{r}(0)$ and calculate $\langle x_i(t) x_k(t)\rangle$ for $t > 0$. We have

$$\langle x_i(t) x_k(t)\rangle = \int_0^t dt_1 \int_0^t dt_2\, \langle v_i(t_1) v_k(t_2)\rangle = \delta_{ik} \int_0^t dt_1 \int_0^t dt_2\, \langle v_i(t_1) v_i(t_2)\rangle = \langle x_i^2(t)\rangle\,\delta_{ik}\,.$$

Here we employ the fact the different components of the velocity are uncorrelated. Now we employ the Langevin method to find the velocity correlation function

$$\langle v_i(t_1) v_i(t_2)\rangle = \langle v_i^2\rangle\, e^{-\lambda|t_1 - t_2|}\,. \tag{8.12}$$

As a result,

$$\begin{aligned}
\langle x_i^2(t)\rangle &= \langle v_i^2\rangle \int_0^t dt_1 \left[\int_0^{t_1} dt_2\, e^{-\lambda(t_1 - t_2)} + \int_{t_2}^t dt_2\, e^{-\lambda(t_2 - t_1)}\right] \\
&= 2\langle v_i^2\rangle \lambda^{-1}\left[t - \lambda^{-1} + e^{-\lambda t}\right]\,.
\end{aligned}$$

At large times, $t \gg \lambda^{-1}$ we obtain (in general)

$$\langle x_i^2(t)\rangle = 2D|t|\,, \quad D \equiv \langle v_i^2\rangle \lambda^{-1}\,. \tag{8.13}$$

The kinetic coefficient D is called the *diffusion constant*. We will discuss its physical meaning later. At small times, $\langle x_i^2(t)\rangle = \langle v_i^2\rangle t^2$ that corresponds to *ballistic* motion.

Onsager relations

The “equations of motion”

$$\dot{x}_i = -\sum_k \lambda_{ik} x_k$$

have a deep internal symmetry. To reveal it let is express the r.h.s. through the thermodynamically conjugated quantities,

$$X_i = -\frac{\partial S}{\partial x_i},$$

which were introduced earlier. We have

$$\langle x_i X_k \rangle = \delta_{ik}, \quad \langle x_i X_k \rangle = (\hat{\gamma}^{-1})_{ik}, \quad \langle X_i X_k \rangle = \gamma_{ik}.$$

If x_i are comparatively small,

$$X_i = \sum_k \gamma_{ik} x_k, \quad \gamma_{ik} = \gamma_{ki}.$$

The latest equation is a consequence of the fact that the coefficients are second derivatives of the entropy. Then we get,

$$\dot{x}_i = -\sum_k \eta_{ik} X_k, \quad \eta_{ik} = \sum_l \lambda_{il} (\hat{\gamma}^{-1})_{lk} = (\hat{\lambda}\hat{\gamma}^{-1})_{ik}.$$

The quantities η_{ik} are called the *kinetic coefficients*, they are just responses to generalized forces. According to the *Onsager principle*,

$$\eta_{ik} = \eta_{ki}.$$

The proof of the principle is based on accurate application of time-reversal symmetry and the property $\langle X_i x_k \rangle = \delta_{ik}$.

If the system is embedded into an external magnetic field **H**, or it rotates as a whole with the angular velocity **Ω** then the time-reversal operation must include reversal of **H** or **Ω**. Then the Onsager principle reads

$$\eta_{ik}(\mathbf{H}) = \eta_{ki}(-\mathbf{H}), \quad \eta_{ik}(\mathbf{\Omega}) = \eta_{ki}(-\mathbf{\Omega}).$$

For proper application of the Onsager principle it is important to find proper thermodynamic forces X_i conjugated to the variables x_i. A convenient way to do that is to study entropy generation. Indeed,

$$\dot{S} = \sum_i \frac{\partial S}{\partial x_i} \dot{x}_i = -\sum_i X_i \dot{x}_i. \tag{8.14}$$

Thus, one can express the entropy generation through the heat release at a given temperature, $\dot{Q}$, as $\dot{S} = T^{-1}\dot{Q}$ and then express it through the quantities x_i in the form (8.14). An example is given below.

Determination of proper generalized forces

According to the general principle, one has to find the entropy production rate, $\dot{S}$, and express is as $-\sum_i \dot{x}_i X_i$. Thus,

$$X_i = -\frac{\partial \dot{S}}{\partial \dot{x}_i}\,.$$

Let us consider an example. An external electric field $\mathbf{E} = -\nabla\varphi$ (where φ is the electrochemical potential) with combination of the temperature gradient create electric current with density $\mathbf{j}$, as well as energy current $\mathbf{w}$. The currents are conventionally expressed as

$$\begin{aligned} \mathbf{j} &= \sigma\mathbf{E} - \eta\nabla T\,, \\ \tilde{\mathbf{w}} \equiv \mathbf{w} - \varphi\mathbf{j} &= \gamma\mathbf{E} - \beta\nabla T\,. \end{aligned} \tag{8.15}$$

Our aim is to determine the corresponding thermodynamic forces, $\mathbf{X_j}$ and $\mathbf{X_w}$. Since the heat release per unit volume consists of the Joule heating $\mathbf{j}\cdot\mathbf{E}$ and of absorption $-\operatorname{div}\mathbf{w}$ and in the linear response approximation $\operatorname{div}\mathbf{w} = \operatorname{div}(\mathbf{w} - \varphi\mathbf{j}) = \operatorname{div}\tilde{\mathbf{w}}$, [3] we have

$$\dot{S} = -\int \frac{\operatorname{div}\tilde{\mathbf{w}}}{T}\,dV + \int \frac{\mathbf{j}\cdot\mathbf{E}}{T}\,dV$$

Let us assume that the total heat flux through the surface is zero. Then the first integral can be calculated by integration by parts, and we have

$$\dot{S} = \int dV\,\left[\tilde{\mathbf{w}}\cdot\nabla\left(\frac{1}{T}\right) + \frac{\mathbf{j}\cdot\mathbf{E}}{T}\right] \equiv -\int dV\,\left[\tilde{\mathbf{w}}\cdot\mathbf{X_w} + \mathbf{j}\cdot\mathbf{X_j}\right]\,.$$

Thus, the generalized forced conjugated to the currents $\mathbf{j}$ and $\mathbf{w}$ are:

$$\mathbf{X_j} = -\mathbf{E}/T\,, \quad \mathbf{X_w} = -\nabla(1/T) = T^{-2}\nabla T\,. \tag{8.16}$$

Compare equations (8.16) and (8.15) we arrive at the following matrix for kinetic coefficients:

$$\hat{\eta} = \begin{pmatrix} -\sigma T & -\eta T^2 \\ -\gamma T & -\beta T^2 \end{pmatrix}.$$

As a result, we have derived an important relation

$$\gamma = \eta T\,.$$

It can be shown that Onsager relations are also valid for the inverse matrix of kinetic coefficients,

$$\hat{\zeta} \equiv (\hat{\lambda}\hat{\gamma}^{-1})^{-1} = \hat{\gamma}\hat{\lambda}^{-1}\,.$$

Exercise:

1. Analyze Onsager relations between the kinetic coefficients given by the equations

$$\begin{aligned} \mathbf{E} &= \rho\mathbf{j} - \alpha\nabla T\,, \\ \mathbf{w} - \varphi\mathbf{j} &= \Pi\mathbf{j} - \varkappa\nabla T\,. \end{aligned} \tag{8.17}$$

2. Express the coefficients ρ, α, Π, $\varkappa$ through the quantities σ, η, γ, β and check the Onsager relations explicitly.

---

[3] Indeed, $\operatorname{div}\varphi\mathbf{j} = \varphi\operatorname{div}\mathbf{j} + \mathbf{j}\cdot\nabla\varphi \approx \varphi\operatorname{div}\mathbf{j} = 0$.

Dissipative function

The problem of dissipation cannot be solved generally within statistical physics. However, a formulation becomes possible if one assumes that the state of the system is fully determined by macroscopic coordinates, Q_i, and by the velocities, $\dot{Q}_i$. That means that high-order derivatives can be neglected. One has also to assume that the velocities are relatively small and high powers of $\dot{Q}_i$ can be neglected, and that the motions consist of small oscillations around some equilibrium positions.

Under that assumptions both kinetic energy, $K(\dot{Q}_i)$, and the potential one, $V(Q_i)$, are *quadratic* functions of their arguments. One can derive the momenta in a usual way,

$$P_i = \frac{\partial K(\dot{Q}_i)}{\partial \dot{Q}_i} .$$

Then we can express $\dot{Q}_i$ in terms of P_i. In the absence of dissipation we come to the usual Hamilton equations of motion,

$$\dot{Q}_i = \frac{\partial K(P_i)}{\partial P_i} , \quad \dot{P}_i = -\frac{\partial V(Q_i)}{\partial Q_i} .$$

Now let us consider the quantities Q_i and P_i as fluctuating quantities x_k. Since the change in the entropy is equal to the minimal work, $W_{\min} = K(P_i) + V(Q_i)$, divided by the temperature of the thermal bath, T, the generalized forces are

$$\begin{aligned} X_{Q_i} &= \frac{1}{T}\frac{\partial W_{\min}}{\partial Q_i} = \frac{1}{T}\frac{\partial V}{\partial Q_i} = -\frac{P_i}{T} , \\ X_{P_i} &= \frac{1}{T}\frac{\partial W_{\min}}{\partial P_i} = \frac{1}{T}\frac{\partial K}{\partial P_i} = \frac{Q_i}{T} , \end{aligned}$$

As a result, $\eta_{Q_i P_i} = -\eta_{P_i Q_i} = -T$. The coefficients satisfy the Onsager relations because one of the quantities (P_i) changes its sign under the time reversal.

Now we can formally write the equations of motion which will allow for dissipative processes. In general, we can add to the r.h.s. of the Hamilton equations any linear combination of X_{P_i}, X_{Q_i} which possess the proper symmetry. However, the equation for $\dot{Q}_i$ should be left unchanged because it is equivalent to the definition of momenta. On the other hand, the additional terms to the equation for $\dot{P}_i$ can contain only X_{P_i}, otherwise the symmetry of kinetic coefficients will be violated. In this way we can write

$$\dot{P}_i = -\frac{\partial V(Q_i)}{\partial Q_i} - \sum_k \eta_{ik} \frac{\partial K(P_k)}{\partial P_k} = -\frac{\partial V(Q_i)}{\partial Q_i} - \sum_k \eta_{ik} \dot{Q}_k , \quad \eta_{ik} = \eta_{ki} .$$

Now we can construct a quadratic form.

$$\mathcal{D} = \frac{1}{2} \sum_{ik} \eta_{ik} \dot{Q}_i \dot{Q}_k$$

which is called the *dissipative function*. The dissipative forces lead to the equation of motion

$$\dot{P}_i = -\frac{\partial V(Q_i)}{\partial Q_i} - \frac{\partial \mathcal{D}(\dot{Q}_i)}{\partial \dot{Q}_i}\,.$$

The energy conservation law has to modified as

$$\begin{aligned}
\frac{d}{dt}(K+V) &= \sum_i \left[\frac{\partial K}{\partial P_i}\dot{P}_i + \frac{\partial V}{\partial Q_i}\dot{Q}_i\right] = \sum_i \dot{Q}_i \left[\dot{P}_i + \frac{\partial V}{\partial Q_i}\right] \\
&= -\sum_i \dot{Q}_i \frac{\partial \mathcal{D}}{\partial \dot{Q}_i} = -2\mathcal{D}\,.
\end{aligned}$$

Note that we applied the symmetry principle which has to be modified in an external magnetic field. For that case out approach is strictly speaking not valid.

8.4 Spectral properties of fluctuations and fluctuation-dissipation theorem

8.4.1 Classical systems

Let us start with a classical system and discuss properties of Fourier transformed fluctuation,

$$x_\omega \equiv \int_{-\infty}^{\infty} x(t)\, e^{i\omega t}\, dt \quad \to \quad x(t) = \int_{-\infty}^{\infty} x_\omega\, e^{-i\omega t}\,\frac{d\omega}{2\pi}\,. \tag{8.18}$$

If $x(t)$ is real, then

$$x_{-\omega} = x^*_\omega\,. \tag{8.19}$$

Since

$$\phi(t'-t) = \int_{-\infty}^{\infty} \frac{d\omega\, d\omega'}{(2\pi)^2} e^{-i(\omega t+\omega' t')}\langle x_\omega x_{\omega'}\rangle\,,$$

and it must be a function of $t'-t$, we have $\langle x_\omega x_{\omega'}\rangle \propto \delta(\omega+\omega')$. The proportionality coefficient is usually denoted as $2\pi(x^2)_\omega$, so

$$\langle x_\omega x_{\omega'}\rangle = 2\pi (x^2)_\omega\, \delta(\omega+\omega')\,. \tag{8.20}$$

Obviously,

$$\begin{aligned}
\phi(t) &= \int \frac{d\omega}{2\pi}\,(x^2)_\omega\, e^{-i\omega t}\,, \\
\langle x^2\rangle &= \phi(0) = \int \frac{d\omega}{2\pi}\,(x^2)_\omega\,.
\end{aligned} \tag{8.21}$$

The quantity $(x^2)_\omega$ is called the *spectral density* of fluctuations. From the known expression (8.8) for $\phi(t)$ we obtain [4]

$$(x^2)_\omega = \frac{2\lambda}{\gamma(\omega^2+\lambda^2)}\,.$$

We can introduce also transform of random forces, y_ω,

$$(y^2)_\omega = (\omega^2+\lambda^2)(x^2)_\omega = 2\lambda/\gamma.$$

Again, the results can be expressed in the form of Langevin approach. For the correlation function of the Fourier components we obtain:

$$\langle y_{i\omega} y_{k\omega'}\rangle = 2\pi (y_i y_k)_\omega \delta(\omega+\omega')\,, \quad (y_i y_k)_\omega = \eta_{ik}+\eta_{ki}\,. \tag{8.22}$$

Here $\hat{\gamma}$ is the matrix of kinetic coefficients.

Example: Fluctuations in one-dimensional oscillator.

Consider one-dimensional oscillator which is at rest is in in the equilibrium position $Q=0$, but able to perform small oscillations in some macroscopic coordinate Q. We write the potential energy as $U=(1/2)m\omega_0^2 Q^2$ where m is the "mass" while ω_0 is the oscillator eigenfrequency. The generalized moment is then $P=m\dot{Q}$. Since the minimal work against the "spring" is equal to the potential energy, $W_{\min}=(1/2)m\omega_0^2Q^2$,

$$f(Q) \propto e^{-W_{\min}/T} = e^{-m\omega_0^2Q^2/2T}\,, \quad \Rightarrow \quad \langle Q^2\rangle = T/m\omega_0^2\,. \tag{8.23}$$

If friction is present, the equations of motion acquire the form

$$\begin{aligned} \dot{Q} &= P/m\,, && (8.24)\\ \dot{P} &= -m\omega_0^2 Q - \gamma P/m\,, && (8.25) \end{aligned}$$

where $-\gamma P/m = -\gamma\dot{Q}$ is the friction force. Now let us take Q and P as random quantities and determine the conjugated thermodynamic forces, X_Q and X_P. They are determined by the entropy generation, i.e. by the minimum work to bring the system from the equilibrium to the state with given Q and P. We have

$$W_{\min} = P^2/2m + m\omega_0^2Q^2/2\,.$$

Consequently,

$$\begin{aligned} X_Q &= T^{-1}\partial W_{\min}/\partial Q = T^{-1}\partial U/\partial Q = m\omega_0^2 Q/T\,, && (8.26)\\ X_P &= T^{-1}\partial W_{\min}/\partial P = T^{-1}\partial K/\partial Q = P/m\,. && (8.27) \end{aligned}$$

Combining Eqs. (8.25) and (8.27) we get the $\hat{\eta}$-matrix has the form

$$\hat{\eta} = \begin{pmatrix} 0 & -T \\ T & \gamma T \end{pmatrix}. \tag{8.28}$$

---

[4] Note that according to general principles of quasistationary fluctuations this expression is valid for ω less that the inverse time for partial equilibrium to be established.

In order to apply these equations for fluctuations we rewrite Eq. (8.25) as

$$\dot{P} = -m\omega_0^2 Q - \gamma P/m + y\,. \tag{8.29}$$

Equation (8.24) is just a definition of the momentum must remain unchanged. According to Eq. (8.22),

$$(y^2)_\omega = 2\eta_{22} = 2\gamma T\,. \tag{8.30}$$

To calculate the required fluctuation spectrum $(Q^2)_\omega$ we substitute $P = m\dot{Q}$ in (8.22), obtaining

$$m\ddot{Q} + \gamma\dot{Q} + m\omega_0^2 Q = y\,. \tag{8.31}$$

Multiplying by $e^{-\omega t}$ and integrating over time, we find

$$(-m\omega^2 - i\omega\gamma + m\omega_0^2)Q_\omega = y_\omega\,,$$

and finally,

$$(Q^2)_\omega = \frac{2\gamma T}{m^2(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2}\,. \tag{8.32}$$

8.4.2 Quantum systems

Now let us turn to a quantum system. First, we have to remind some basic concepts of quantum mechanics to be able consider time-dependent quantities.

Revisiting quantum mechanics

According to quantum mechanics, the state of the system can be described by a time-dependent wave function $\Psi(t)$ satisfying the Schrödinder equation,

$$i\hbar(\partial\Psi/\partial t) = \mathcal{H}\Psi \tag{8.33}$$

where $\mathcal{H}$ is the Hamiltonian. The states ψ_n which satisfy the *stationary* Schrödinder equation,

$$\mathcal{H}\psi_n(q) = E_n\psi_n(q)\,, \tag{8.34}$$

i.e. *eigenstates* of the Hamiltonian are called the *stationary states*. We will also use Dirac notations,

$$\psi_n(q) \equiv |n\rangle,\, \psi^*(q) \equiv \langle n|\,.$$

Since eigenfunctions of the Schrödinder equation form a complete set,

$$\int dq\,\psi_n^*(q)\psi_n(q) = 1\,. \tag{8.35}$$

For a stationary state,

$$\Psi_n(t) = e^{-(i/\hbar)E_n t}\,\psi_n(q)\,. \tag{8.36}$$

In general,

$$\Psi(t) = \sum_n a_n \Psi_n(t) = \sum_n a_n e^{-(i/\hbar)E_n t} \psi_n(q) . \tag{8.37}$$

The mean value of any physical quantity, $\bar{f}$, represented by the operator $\hat{f}$ is then

$$\bar{f}(t) = \sum_{nm} a_n^* a_m f_{nm}(t) \tag{8.38}$$

where

$$f_{nm}(t) = \int \Psi_n^*(t) \hat{f} \Psi_m(t) . \tag{8.39}$$

Now, let us assume that the operator $\hat{f}$ does not contain time explicitly. Then using Eq. (8.36) we have,

$$f_{nm}(t) = f_{nm} e^{i\omega_{nm} t} , \tag{8.40}$$

where

$$f_{nm} \equiv \langle n|\hat{f}|m\rangle = \int dq \, \psi_n^* \hat{f} \psi_m(q) , \quad \omega_{nm} \equiv \frac{E_n - E_m}{\hbar} . \tag{8.41}$$

8.4.3 The generalized susceptibility

Let us now assume that the body is subjected to some perturbation,

$$\hat{V}(t) = -f(t)\hat{x} .$$

Then the average value

$$\bar{x}(t) = \int_0^\infty \alpha(\tau) \, f(t-\tau) \, d\tau . \tag{8.42}$$

Consequently,

$$\bar{x}(\omega) = \alpha(\omega) \, f(\omega) , \tag{8.43}$$

where

$$\alpha(\omega) = \int_0^\infty dt \, \alpha(t) \, e^{i\omega t} . \tag{8.44}$$

$\alpha(\omega) \equiv \alpha'(\omega) + i\alpha''(\omega)$ is called the *linear response function.* Since $\bar{x}$ is real

$$\alpha(\omega) = \alpha^*(-\omega) , \quad \to \quad \alpha'(\omega) = \alpha'(-\omega) , \alpha''(\omega) = -\alpha''(-\omega) .$$

For a periodic perturbation,

$$\hat{V} = -\frac{1}{2} \left[f e^{-i\omega t} + f^* e^{i\omega t} \right] \hat{x} ,$$

and

$$\bar{x}(\omega) = \frac{1}{2} \left[\alpha(\omega) f e^{-i\omega t} + \alpha(-\omega) f^* e^{i\omega t} \right] . \tag{8.45}$$

Then in general the dissipation is

$$\frac{dE}{dt} = \overline{\frac{\partial \hat{\mathcal{H}}}{\partial t}} = -\bar{x} \frac{df}{dt} . \tag{8.46}$$

Substituting $\bar{x}$ from Eq. (8.45),

$$f(t) = \frac{1}{2}\left[f\,e^{-i\omega t} + f^* e^{i\omega t}\right] \qquad (8.47)$$

and averaging over time we obtain

$$Q = \left\langle \frac{dE}{dt} \right\rangle_t = \frac{1}{4}(\alpha^* - \alpha)|f|^2 = \frac{1}{2}|f|^2\omega\alpha''(\omega)\,. \qquad (8.48)$$

Analytical properties of generalized susceptibility. Kramers-Kronig relations

Let us regard ω as a complex variable, $\omega = \omega' + i\omega''$ and consider the properties of $\alpha(\omega)$ in the upper half of the ω-plane. From the definition (8.42) it follows that $\alpha(t)$ is finite at any positive t. Consequently, $\alpha(\omega)$ is a one-valued regular function in the upper half of the ω-plane. Indeed, for $\omega'' > 0$ the integrand of Eq. (8.44) contains the factor $e^{-\omega'' t}$ and since $\alpha(t)$ is finite the integral converges. The function $\alpha(\omega)$ has no singularities at the real axis $(\omega'' = 0)$, except possible at the origin. Note that the definition (8.44) is not valid for the lower half of the ω-plane since the integral diverges at $\omega'' < 0$.

The conclusion that $\alpha(\omega)$ is regular in the upper half-plane is the consequence of the *causality principle*. It is because of this principle the integration over time in Eq. (8.42) is taken only over times previous to t. It is also evident from the definition (8.44) that

$$\alpha(-\omega^*) = \alpha^*(\omega)\,. \qquad (8.49)$$

Hence for a pure imaginary frequency, $\omega = i\omega''$, we have $\alpha(i\omega'') = \alpha^*(i\omega'')$, i.e. $\alpha(\omega)$ is real.

The following theorem can be proved: *The function* $\alpha(\omega)$ *does not take real numbers at any finite point of the upper half-plane except on the imaginary axis, where it decreases monotonously from a positive value* $\alpha_0 > 0$ *at* $\omega = i0$ *to zero at* $\omega = i\infty$. In particular, it follows that $\alpha(\omega)$ has no zeros in the upper half-plane. The proof is given in the book [1], §123.

Let us now derive a formula relating the real and imaginary part of $\alpha(\omega)$. To do so, we choose some real positive value ω_0 of ω, integrate the expression $\alpha(\omega)/(\omega - \omega_0)$ round the contour C shown in Fig. 8.1, and then tend the outer contour radius R to infinity and the radii ρ of the inner circles to zero.

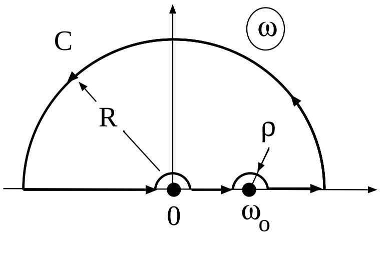

Figure 8.1: On the derivation of analytical properties of the susceptibility.

Since at infinity $\alpha \to 0$, and the ratio $\alpha(\omega)/(\omega - \omega_0)$ tends to zero more rapidly than ω^{-1}. Since the function $\alpha(\omega)$ is regular and possible divergence points $\omega = 0$ and $\omega = \omega_0$ are excluded, the result must vanish,

$$\oint_C d\omega\,\frac{\alpha(\omega)}{\omega - \omega_0} = 0\,.$$

The integral over the infinite semicircle is also zero, so we are left with the integral over the real axis and the parts around the $\omega = 0$ and $\omega = \omega_0$. The first part in not important since $\alpha(0)$ is finite. The integration around ω_0 yields $-i\pi\alpha(\omega_0)$ since the direction is clockwise. Consequently,

$$\lim_{\rho\to 0}\left\{\int_{-\infty}^{\omega_0-\rho} d\omega\,\frac{\alpha(\omega)}{\omega-\omega_0} + \int_{\omega_0+\rho}^{\infty} d\omega\,\frac{\alpha(\omega)}{\omega-\omega_0}\right\} - i\pi\alpha(\omega_0) = 0\,.$$

The first item is just the principle value of the integral, so we obtain

$$i\pi\alpha(\omega_0) = P\int_{-\infty}^{\infty} d\omega\,\frac{\alpha(\omega)}{\omega-\omega_0}\,. \tag{8.50}$$

Here the variable of integration, ω has only real values. For convenience we replace the notations as

$$\omega \to \xi\,, \quad \omega_0 \to \omega\,,$$

to obtain the famous Kronig-Kramers relations (1927):

$$\alpha'(\omega) = \frac{1}{\pi}P\int_{-\infty}^{\infty} d\xi\,\frac{\alpha''(\xi)}{\xi-\omega}\,, \tag{8.51}$$

$$\alpha''(\omega) = -\frac{1}{\pi}P\int_{-\infty}^{\infty} d\xi\,\frac{\alpha'(\xi)}{\xi-\omega}\,. \tag{8.52}$$

Since $\alpha''(\omega)$ is the odd function, Eq. (8.51) can be rewritten as

$$\alpha'(\omega) = \frac{2}{\pi}P\int_{0}^{\infty} d\xi\,\frac{\xi\alpha''(\xi)}{\xi^2-\omega^2}\,. \tag{8.53}$$

If the function $\alpha(\omega)$ has a pole at $\omega = 0$, near which $\alpha(\omega) = iA/\omega$, then we get instead of Eq. (8.52):

$$\alpha''(\omega) = -\frac{1}{\pi}P\int_{-\infty}^{\infty} d\xi\,\frac{\alpha'(\xi)}{\xi-\omega} + \frac{A}{\omega}\,. \tag{8.54}$$

The Kramers-Kronig relations are very important since they allow to reconstruct the whole response function from its real or imaginary part. This is a consequence of the causality principle.

Quantum expression for susceptibility

On the other hand,the probability for the transition $n \longrightarrow m$ according to the Fermi golden rule is

$$w_{mn} = \frac{2\pi}{\hbar}\frac{|f|^2}{4\hbar}|x_{mn}|^2\left[\delta(\omega+\omega_{nm}) + \delta(\omega+\omega_{mn})\right]\,. \tag{8.55}$$

The total absorbed power due to all transitions from the state n is then

$$Q_n = \sum_m \hbar\omega_{mn}w_{mn} = \frac{\pi\omega|f|^2}{2\hbar}\sum_m |x_{nm}|^2\left[\delta(\omega+\omega_{nm}) - \delta(\omega+\omega_{mn})\right]\,. \tag{8.56}$$

Comparing Eqs. (8.48) and (8.56) we write down the response function[5] α'' for n-th state as

$$\alpha''_n(\omega) = \frac{\pi}{\hbar}\sum_m |x_{nm}|^2 \left[\delta(\omega+\omega_{nm}) - \delta(\omega+\omega_{mn})\right] . \tag{8.57}$$

This quantity should be averaged over the Gibbs distribution of the initial states occupations, $f_n = Z^{-1}e^{-\beta E_n}$, as

$$\alpha''(\omega) = \mathrm{Tr}\,\hat{\rho}\hat{\alpha} = \sum_n f_n \alpha''_n(\omega) \tag{8.58}$$

where f_n are diagonal elements of the density matrix $\hat{\rho} = Z^{-1}e^{\beta\mathcal{H}}$, see Eq. (7.12). We have

$$\alpha''(\omega) = \frac{\pi}{\hbar}\sum_{nm}(f_n - f_m)\,|x_{nm}|^2\,\delta(\omega+\omega_{nm})\,. \tag{8.59}$$

Immediately, we get

$$\alpha'(\omega) = \frac{1}{\hbar}\sum_{nm}|x_{nm}|^2\,\frac{f_n - f_m}{\omega_{mn}-\omega}\,, \tag{8.60}$$

or, in the complex form,

$$\alpha(\omega) = \frac{1}{\hbar}\sum_{nm}|x_{nm}|^2\,\frac{f_n - f_m}{\omega_{mn}-\omega-i0}\,, \tag{8.61}$$

8.4.4 Fluctuation-dissipation theorem: Proof

According to quantum mechanics, one has to define a real quantity which is an average of a Hermitian operator,

$$\frac{1}{2}\left\{\hat{x}_\omega,\hat{x}_{\omega'}\right\} \equiv \frac{1}{2}\left\{\hat{x}_\omega\hat{x}_{\omega'} + \hat{x}_{\omega'}\hat{x}_\omega\right\} .$$

Obviously, the correlation function must depend only on the time difference, and one can write down the spectral correlator as

$$\frac{1}{2}\langle \hat{x}_\omega\hat{x}_{\omega'} + \hat{x}_{\omega'}\hat{x}_\omega\rangle \equiv 2\pi (x^2)_\omega\,\delta(\omega+\omega')\,. \tag{8.62}$$

This relation can be regarded as a *definition* for $(x^2)_\omega$. Using the definition, as well as the fact that $(x^2)_\omega$ is real and is an even function of ω we get

$$\phi(t) = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,x^2)_\omega e^{-i\omega t} \quad \to \quad (x^2_\omega = \int_{-\infty}^{\infty} dt\,\phi(t)\,e^{i\omega t}\,.$$

In particular,

$$\phi(0) = \langle x^2\rangle = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,(x^2)_\omega\,.$$

---

[5] α'' is the imaginary part of the *susceptibility*.

The aim of the present section is to relate the spectral correlation function $(x^2)_\omega$ to the dissipation in the system. Let us consider a quantum body *in a given state*, say n. Then, the average value to the correlator is its diagonal matrix element,

$$\frac{1}{2}\langle \hat{x}_\omega \hat{x}_{\omega'} + \hat{x}_{\omega'}\hat{x}_\omega \rangle_{nn} = \frac{1}{2}\sum_m \left[(x_\omega)_{nm}(x_{\omega'})_{mn} + (x_{\omega'})_{nm}(x_\omega)_{mn}\right] .$$

Here $(x_\omega)_{nm}$ is the Fourier transform of the nm matrix element of the operator $\hat{x}(t)$,

$$(x_\omega)_{nm} = \int_{-\infty}^{\infty} dt\, x_{nm}(t)\, e^{i\omega t} = \int_{-\infty}^{\infty} dt\, x_{nm}\, e^{i(\omega_{nm}+\omega)t} = 2\pi x_{nm}\, \delta(\omega_{nm} - \omega) .$$

Since x is real $x_{nm} = x^*_{mn}$ and

$$\delta(\omega_{nm} + \omega)\delta(\omega_{mn} + \omega') = \delta(\omega_{nm} + \omega)\delta(\omega + \omega')$$

we can extract $2\pi\delta(\omega + \omega')$ at get:

$$(x^2)_\omega = \pi \sum_m |x_{nm}|^2 \left[\delta(\omega + \omega_{nm}) + \delta(\omega + \omega_{mn})\right] . \tag{8.63}$$

Again, the expression for the fluctuations should be averaged over the Gibbs distribution of the initial state occupations, $f_n = Z^{-1}e^{-\beta E_n}$. Averaging, we obtain

$$\begin{aligned}
(x^2)_\omega &= \pi \sum_{nm} f_n\, |x_{nm}|^2 \left[\delta(\omega + \omega_{nm}) + \delta(\omega + \omega_{mn})\right] \\
&= \pi \sum_{nm} (f_n + f_m) |x_{nm}|^2\, \delta(\omega + \omega_{nm}) = \pi \sum_{nm} f_n \left(1 + e^{\beta\hbar\omega_{nm}}\right) |x_{nm}|^2\, \delta(\omega + \omega_{nm}) \\
&= \pi \left(1 + e^{-\beta\hbar\omega}\right) \sum_{nm} f_n\, |x_{nm}|^2\, \delta(\omega + \omega_{nm}) .
\end{aligned} \tag{8.64}$$

Now, let us note that Eq. (8.59) can be rewritten in the form

$$\alpha''(\omega) = \frac{\pi}{\hbar}\left(1 - e^{-\beta\hbar\omega}\right) \sum_{nm} f_n\, |x_{nm}|^2\, \delta(\omega + \omega_{nm}) . \tag{8.65}$$

Comparing Eqs. (8.64) and (8.65), we obtain

$$(x^2)_\omega = \hbar\alpha''(\omega)\coth(\beta\hbar\omega/2) \quad \to \quad \langle x^2 \rangle = \frac{\hbar}{\pi}\int \alpha''(\omega)\coth(\beta\hbar\omega/2)\, d\omega . \tag{8.66}$$

This is the famous fluctuation-dissipation theorem (Callen & Welton, 1951).

For small frequencies, when $\hbar\omega \ll T$, one can approximate $\coth(\beta\hbar\omega/2)$ as

$$\coth(\beta\hbar\omega/2) \approx 2T/\hbar\omega .$$

As a result, instead of (8.66) we get

$$(x^2)_\omega = T\,\frac{\alpha''(\omega)}{\omega} . \tag{8.67}$$

Exercise: *Check that fluctuations in the one-dimensional classical oscillator (example in p. 102) given by Eq. (8.32) obey the classical limit (8.66) of the fluctuation-dissipation theorem.*

An important example is fluctuations of current density in a resistor. In this case $x_i = j_i$, $\alpha'' = \omega \Re \sigma(\omega)$, where $\sigma(\omega)$ is the complex conductivity. Thus

$$(j^2)_\omega = \hbar\omega \Re \sigma(\omega) \coth\left(\frac{\hbar\omega}{2T}\right) .$$

Note that all the concept is valid only for small fluctuations near the equilibrium state.

Exercise: *1) Find spectrum of fluctuations of the current I through the circuit shown in Fig. 8.2, $(I^2)_\omega$, for a fixed voltage V_{12} between the points 1 and 2) Calculate the integrated spectrum, $\int (I^2)_\omega \, (d\omega/2\pi)$ and compare it with the square of the average current, $\langle I \rangle^2$.*

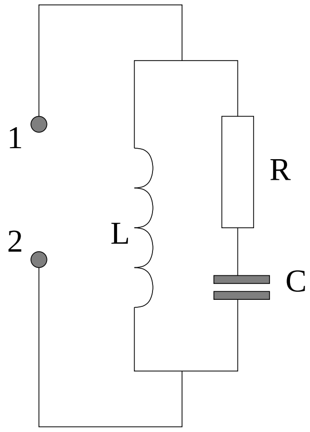

Figure 8.2: On the current fluctuations. L is the inductance, C is the capacitance, while R is the resistance. The circuit is kept under constant voltage V_{12}.

Chapter 9

Stochastic Processes

Now we discuss other aspects of random processes which take place in course of evolution of nonequilibrium states to equilibrium.

9.1 Random walks

The simplest stochastic process is random walk in one dimension. A particle can make a step to the right with the probability p or to the left with probability $q = 1 - p$.

We are interested to find the displacement S after $N \gg 1$ steps. It can vary between $+N$ and $-N$.

The probability of the walk with $S = +N$ is $P_N(N) = p^N$. If there occurs 1 step to the left the displacement is $S = N - 2$, the probability being $P_N(N-1) = Np^{N-1}q$ because the step to the left can occur at any of the N different times (see Fig. 9.1). In general, the probability to find a walk with R steps to the right from total N steps is

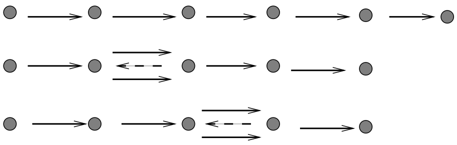

Figure 9.1: On the 1D random walk.

$$P_N(R) = \frac{N!}{R!(N-R)!}\, p^R q^{N-R}\,. \tag{9.1}$$

This distribution is called *binomial* or *Bernoulli*. Certainly

$$\sum_{R=0}^{N} P_N(R) = (p+q)^N = 1\,.$$

As a result, we can easily calculate the average displacement

$$\begin{aligned}
\langle S\rangle &= \langle R\rangle-(N-\langle R\rangle)=2\langle R\rangle-N \\
&= -N+2\sum_R \frac{N!}{R!(N-R)!}Rp^Rq^{N-R} \\
&= -N+2p\frac{\partial}{\partial p}\sum_R \frac{N!}{R!(N-R)!}p^Rq^{N-R} \\
&= -N+2\frac{\partial}{\partial p}(p+q)^N=N(p-q)\,.
\end{aligned}$$

The symmetric random walk can be generated by tossing a coin. One makes N attempts. To obtain the statistical average one has to make $M \gg 1$ runs. This set forms a statistical ensemble, the average being

$$\langle S\rangle=\frac{1}{M}\sum_{m=1}^{M}S_m\,.$$

In general case the way is more difficult, one needs a random number generator.

One can also obtain

$$\langle R^2\rangle=\left(p\frac{\partial}{\partial p}\right)^2(p+q)^N=(Np)^2+Npq$$

to obtain

$$\langle S^2\rangle=4\langle R^2\rangle-4N\langle R\rangle+N^2=N^2(p-q)^2+4Npq\,.$$

Thus

$$\langle(\Delta S)^2\rangle=\langle S^2\rangle-\langle S\rangle^2=4Npq \quad\to\quad =N \quad (\text{for } p=q=1/2)\,.$$

For very large N the Bernoulli distribution $P_N(R)=(N!/R!(N-R)!)p^R(1-p)^{N-R}$ can be approximated by the Gaussian distribution. Using the Stirling's formula for $n!$ we obtain

$$P_N(R)=\frac{1}{\sqrt{2\pi Npq}}e^{-(R-Np)^2/2Npq}\,.$$

Since $R=(N+S)/2$ we have

$$P_N(S)=\frac{1}{\sqrt{8\pi Npq}}e^{-[S-N(2p-1)]^2/8Np(1-p)}=\frac{1}{\sqrt{2\pi}\sigma}e^{-(S-\bar{S})^2/2\sigma^2}\,.$$

Here

$$\bar{S}=N(p-q)=N(2p-1)\,,\quad \sigma=\sqrt{\langle(\Delta S)^2\rangle}=4Np(1-p)\,.$$

If the number of steps is very large then we can think that each step is very small and come to continuous description. Assuming $x=Sa$ and regarding the time step $t=N\tau$. Then we obtain for the symmetric case the probability distribution

$$P(x,t)=\sqrt{\frac{\tau}{2\pi a^2t}}\,e^{-a^2\tau/2a^2t}=\frac{1}{\sqrt{4\pi Dt}}e^{-x^2/4Dt}\,,\quad D\equiv\frac{a^2}{2\tau}\,.$$

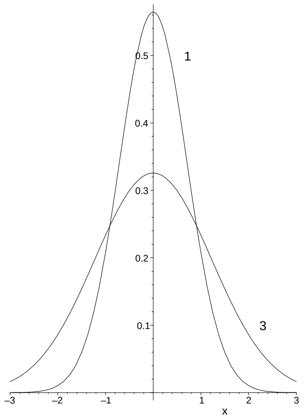

Figure 9.2: $P(x,t)$ as a function of x at $Dt = 1$ and 3.

This function is shown in Fig. 9.2 We immediately obtain

$$\int_{-\infty}^{\infty} P(x,t) = 1, \quad \langle x \rangle = 0, \quad \langle x^2 \rangle = 2Dt \,.$$

The above expressions can be easily generalized for multidimensional space keeping in mind that the probability is multiplicative. For 3D case we obtain,

$$P(r,t) = \frac{1}{(4\pi Dt)^{3/2}} e^{-r^2/4Dt} \,, \quad D = \frac{a^2}{6\tau} \,.$$

We have

$$\langle x_i(t) x_j(t) \rangle = 2\delta_{ij} D|t| \,, \quad \langle r^2 \rangle = 3\langle x^2 \rangle = 6DT \,.$$

Now we can map the above results on the results on Brownian motion obtained in Sec. 8.3 by the Langevin method. Indeed, we can assume that during the time $\lesssim \lambda^{-1}$ between collisions the displacement is $a \sim v/\lambda$. Thus,

$$D = \langle v_i^2 \rangle \lambda \sim a^2/\lambda \,.$$

9.1.1 Relation to diffusion

Consider a drop of colored dye in the water. The concentration of dye obeys the continuity equation,

$$\frac{\partial C}{\partial t} + \operatorname{div} \mathbf{J} = 0$$

where $\mathbf{J}$ is the current of colored particles. It is driven by gradients, $\mathbf{J} = -D \operatorname{grad} C$. In this way we get the *diffusion* equation

$$\frac{\partial C}{\partial t} - D\nabla^2 C = 0 \,. \tag{9.2}$$

Let us consider the initial condition

$$C(\mathbf{r},0) = N\,\delta(\mathbf{r}) \,.$$

Here N is the total number of particles in the drop. The solution with the above initial condition can be easily obtained by the Fourier method. Since

$$C(\mathbf{r},t)=\int\frac{d^3k}{(2\pi)^3}c(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}}\,,\quad\text{and}\quad c(\mathbf{k},t)=Ne^{-Dk^2t}$$

we obtain

$$C(\mathbf{r},t)=\int N\frac{d^3k}{(2\pi)^3}e^{-Dk^2t+i\mathbf{k}\cdot\mathbf{r}}=\frac{N}{(4\pi Dt)^{3/2}}e^{-r^2/4Dt}\,.$$

To derive this formula it is convenient to use spherical coordinates and to direct the polar axis along $\mathbf{r}$. The integral over the angles is

$$\int_0^\pi\sin\theta\,d\theta\int_0^{2\pi}d\phi e^{-kr\cos\theta}=2\pi\cdot2\int_0^1d\mu\cos kr\mu=4\pi\frac{\sin kr}{kr}\,.$$

Then,

$$\frac{4\pi}{8\pi^3}\int_0^\infty k^2dk\frac{\sin kr}{kr}e^{-Dk^2t}=\frac{1}{(4\pi Dt)^{3/2}}e^{-r^2/4Dt}\,.$$

The relation between diffusion and random walk has been realized by A. Einstein in 1905. His reasoning is like that.

Consider a thin layer between x and $x+\Delta x$. Let the number of particles in this layer at time t will be denoted as $N(x,t)$. If we introduce the probability $P(a,\tau)\,\Delta a$, to move a distance falling within a small interval between a and $a+\Delta a$ during a short time interval τ, we can write

$$\Delta N(x,t+\tau)=P(a,\tau)\,\Delta a\,\Delta N(x-a)\,.$$

Since $C=\Delta N/\Delta x$ we get an integral equation:

$$C(x,t+\tau)=\int_{-\infty}^{\infty}da\,P(a,\tau)\,C(x-a,t)\,. \tag{9.3}$$

Since a and τ are small one can expand the l.h.s. in powers of τ, while the r.h.s. in powers of a. Using the equalities

$$\int da\,P(a,\tau)=1\,,\qquad\int da\,aP(a,\tau)=0\quad\text{(we assume symmetric distribution)}$$

and the definition

$$\frac{1}{2\tau}\int da\,a^2P(a,\tau)=\langle x^2(\tau)\rangle/2\tau\equiv D\,,$$

we arrive at the same 1D diffusion equation as for the Brownian motion (with correct coefficients).

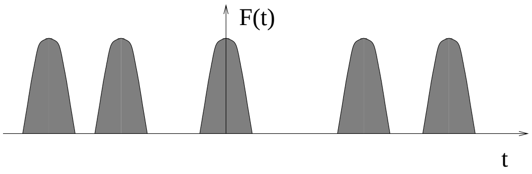

Figure 9.3: On the random pulse process.

9.2 Random pulses and shot noise

Random pulses, or pulsed random processes are encountered very often. In the simplest case the pulses have the same shape, and the only random quantity is the instant at which a pulse starts, Fig. 9.3. One well known example is the current in the anode circuit of the thermionic tube. Each electron escapes the cathode and moves in the space charge region toward the anode producing anode current $i(t)$,

$$\int dt\, i(t) = e\,.$$

There are many examples of this type of process, e. g. impacts of the gas molecules on the wall of the container. Let us denote the number of pulse as k, t_k being its starting time. Then the random quantity can be expressed as

$$x(t) = \sum_k F(t - t_k; \mathbf{a}_k)\,.$$

It is usually assume that

1. The random quantities t_k and $\mathbf{a}_k$ are *statistically independent* for different k, and their distribution functions are k-independent.

2. The probability of the time t_k being in the range between t and $t + dt$ is constant and equal to $\nu\, dt$.

Below we shall use the above example to introduce the important concept - characteristic function of the random quantity. For a quantity x it is defined as

$$\phi_x(u) \equiv \langle e^{iux} \rangle = \int_{-\infty}^{\infty} dx\, p(x)\, e^{iux}\,.$$

Here $p(x)$ is the probability distribution for x while the average is performed over realizations of the random process. For our case

$$\phi_{x_k}(u) = \int d\mathbf{a}\, w_{\mathbf{a}}(\mathbf{a}) \left\{ \frac{1}{t_m} \int_{-t_m/2}^{t_m/2} dt\, \exp[iuF(t - t_k; \mathbf{a})] \right\}.$$

Here $w_{\mathbf{a}}(\mathbf{a})$ is the distribution of the parameters $\mathbf{a}$, while t_m is the measuring time interval. If n pulses took place during the time t_m we obtain

$$\phi_x(u|n) = [\phi_{x_k}(u)]^n$$

which then must be averaged over the Poisson distribution,

$$P_n = \frac{\bar{n}^n}{n!} e^{-\bar{n}}, \quad \bar{n} = \nu t_m .$$

Since

$$e^{-\bar{n}} \sum_n \frac{[\bar{n}\phi_{x_k}(u)]^n}{n!} = \exp\{\bar{n}[\phi_{x_k}(u) - 1]\} = \exp\{\nu t_m [\phi_{x_k}(u) - 1]\}$$

we obtain for the well separated pulses

$$\phi_x(u) = \exp\left\{\nu \int d\mathbf{a}\, w_{\mathbf{a}}(\mathbf{a}) \int_{-\infty}^{\infty} dt\, [\exp(iuF(t,\mathbf{a})) - 1]\right\} .$$

Having the characteristic function we can calculate any moment of x:

$$\langle x \rangle = \left(\frac{\partial \phi_x(u)}{\partial iu}\right)_{u=0} = \nu \int d\mathbf{a}\, w_{\mathbf{a}}(\mathbf{a}) \int_{-\infty}^{\infty} dt\, F(t,\mathbf{a}) , \tag{9.4}$$

$$\langle (\delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 = \left(\frac{\partial^2 \ln \phi_x(u)}{\partial (iu)^2}\right)_{u=0} = \nu \int d\mathbf{a}\, w_{\mathbf{a}}(\mathbf{a}) \int_{-\infty}^{\infty} dt\, F^2(t,\mathbf{a}) , \tag{9.5}$$

$$\langle \delta x(t_1) \delta x(0) \rangle = \nu \int d\mathbf{a}\, w_{\mathbf{a}}(\mathbf{a}) \int_{-\infty}^{\infty} dt\, F(t;\mathbf{a}) F(t+t_1;\mathbf{a}) . \tag{9.6}$$

The noise is conventionally measured as doubled Fourier component of the correlation function (9.6). We obtain (check!):

$$S_x(\omega) = 2\nu \int d\mathbf{a}\, w_{\mathbf{a}}(\mathbf{a})\, |F(\omega;\mathbf{a})|^2 .$$

Since $F(0,\mathbf{a}) \equiv q$ where q is the charge of the particle, we get a very universal formula

$$S_x(0) = 2q^2\nu \quad \to \quad 2eI$$

where I is the electric current in the case of thermionic tube (Schottky, 1918).

9.3 Markov processes

Many stochastic processes belong to the class of *Markov*, or Markovian processes. To introduce a definition let us consider a random function of time, $x(t)$ which is *known* at n instants $t_1 < t_2 < \ldots < t_n$. Let the time intervals be *much greater* than the microscopic time scale. We can characterize the random process by the probability

$$w_n(x_1,t_1;x_2,t_2;\ldots;x_n,t_n)\, dx_1 \cdots dx_n$$

that the random quantity at t_i lies within the range $(x_i, x_i + dx_i)$. We can now relate w_n to w_{n-1} by the conditional probability $P_n(x_n,t_n|x_1,t_1;x_2,t_2;\ldots;x_{n-1},t_{n-1})$ as

$$\begin{aligned} w_n(x_1,t_1;x_2,t_2;\ldots;x_n,t_n) &= w_{n-1}(x_1,t_1;x_2,t_2;\ldots;x_{n-1},t_{n-1}) \\ &\quad \times P_n(x_n,t_n|x_1,t_1;x_2,t_2;\ldots;x_{n-1},t_{n-1}) . \end{aligned}$$

The process is called the Markov one if the conditional probability depends *only on the value* x_{n-1} *at the last instant* t_{n-1} *which precedes* t_n,

$$P_n(x_n,t_n|x_1,t_1;x_2,t_2;\ldots;x_{n-1},t_{n-1}) = P_2(x_n,t_n|x_{n-1},t_{n-1}) \equiv P(x_n,t_n|x_{n-1},t_{n-1})\,. \tag{9.7}$$

It means that the system forgets all but the last previous values of the random process. Both random walks and pulse processes are the Markov ones.

Let us consider a Markov process where the random variable $\xi(t)$ acquires discrete value $\{x_k\}$. Then the conditional probability can be uncoupled as

$$P(x_k,t|x_i,t_0) = \sum_j P(x_k,t|x_j,\theta)P(x_j,\theta|x_i,t_0)\,, \quad t_0 < \theta < t\,. \tag{9.8}$$

This is the *Markov equation*. For a *homogeneous* random quantity $\xi(t)$ depending in discrete times t_n the above equation acquires the form

$$P(x_k|n-l,0) = \sum_j P(x_k|x_j,n-m)P(x_j|m-l,x_i)\,, \quad l < m < n\,. \tag{9.9}$$

9.4 Discrete Markov processes. Master equation

There is an important class of the processes where the random quantity acquires only discrete values $x_i, (i = 1,2,\ldots,N)$. An example is the *random telegraph noise* produced by a defect hopping between two states in the lattice, see Fig. 9.4.

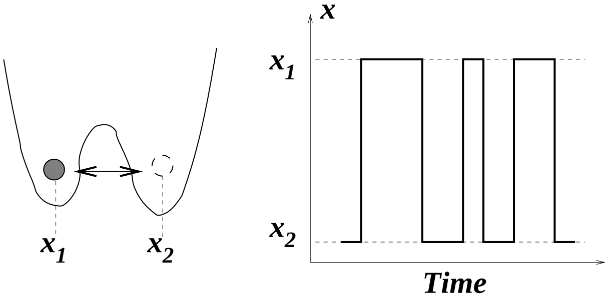

Figure 9.4: Random telegraph noise in the defect position.

In the discrete case one can introduce the probability w_i to find the system in the i-th state, as well as transition probability p_{ik}. For two-state system we can denote, $w_1 = w,\ p_{12} = p,\ p_{21} = q$. Since $w_2 = 1 - w_1 = 1 - w$ we can write the *master equation* for the probabilities in the form

$$\frac{dw}{dt} = -pw + q(1-w)\,. \tag{9.10}$$

The stationary solution of this equation is

$$w^0 = \frac{q}{p+q}\,, \quad w_2^0 = 1 - w^0 = \frac{q}{p+q}\,.$$

Consequently,

$$\langle x\rangle = \frac{qx_1+px_2}{p+q}, \quad \langle(\delta x)^2\rangle = \frac{pq(x_2-x_1)^2}{(p+q)^2}.$$

We can also obtain the correlation function $\langle \delta x(t)\delta x(0)\rangle$. To calculate the correlation function one has to find conditional probability, $P(x,t|x_1,t_1)$. Obviously, at $t \ge t_1$ it has the form

$$P(x,t|x_1,t_1) = w(x) + [\delta(x-x_1) - w(x)]g(t,t_1)$$

where $w(x)$ is the stationary probability to find the system in the state x, while $g(t,t_1)$ is the probability that the state remains *unchanged* during the time between t_1 and $t \ge t_1$. The quantity $g(t,t_1)$ satisfies Eq. (9.10) with the initial condition

$$g(t,t_1)|_{t=t_1+0} = 1. \tag{9.11}$$

The physical meaning of $g(t,t_1)$ is the probability It is natural to look for a non-stationary solution in the form

$$g(t,t_1) = g_0\, e^{-\lambda|t-t_1|}.$$

Substituting that into the master equation (9.10) we get $\lambda = p+q$, and from the initial condition (9.11) $g_0 = 1$.

$$g(t,t_1) = e^{-(p+q)|t-t_1|}.$$

Since at $t-t_1 \gg (p+q)^{-1}$ we have $\langle x(t)x(t_1)\rangle = \langle x\rangle^2$, or $\langle(\delta x)^2\rangle = 0$, we obtain

$$\langle \delta x(t)\delta x(0)\rangle = \langle(\delta x)^2\rangle\, e^{-(p+q)|t|} = \frac{pq(x_1-x_2)^2}{(p+q)^2}\, e^{-(p+q)|t|}.$$

Consequently the noise (doubled Fourier transform) is

$$S(\omega) = (x_1-x_2)^2 \frac{2pq}{(p+q)^2}\cdot\frac{1}{\pi}\frac{p+q}{(p+q)^2+\omega^2}.$$

Thus noise spectrum is a Lorenzian. The integral noise is

$$\int_{-\infty}^{\infty} d\omega S(\omega) = (x_1-x_2)^2\frac{2pq}{(p+q)^2} = 2\langle(\delta x)^2\rangle.$$

Let us for simplicity assume that $p = q = \lambda/2$, than

$$S(\omega) = \frac{(x_1-x_2)^2}{2\pi}\,\frac{\lambda}{\lambda^2+\omega^2}.$$

In many random systems the fluctuation are induced by structural defects with different transition rates λ. Usually *logarithm* of λ is distributed almost uniformly in a wide region between $\lambda_{\min}$ and $\lambda_{\max}$, and for realistic values of ω the inequalities

$$\lambda_{\max}^{-1} \ll \omega \ll \lambda_{\min}^{-1} \tag{9.12}$$

hold. Then the distribution of λ is

$$w(\lambda)\,d\lambda = \frac{n_d}{\mathcal{L}}\frac{d\lambda}{\lambda}, \quad \mathcal{L} \equiv \ln(\lambda_{\max}/\lambda_{\min})\,. \tag{9.13}$$

Here n_d is the density of the defects. An example leading to the distribution (9.13) is so-called tunneling defects formed by an atom which can tunnel between two adjacent states in the material. Since logarithm the tunneling probability is proportional to the tunneling distance and the latter is uniformly distributed we arrive at the distribution (9.13).

Averaging the noise over the distribution (9.13) we get

$$S(\omega) = \frac{n_d(x_1-x_2)^2}{2\mathcal{L}}\frac{1}{\omega} \tag{9.14}$$

provided ω satisfies inequalities (9.12). Noise with the spectrum (9.14) is usually observed at low-frequencies, it is called the *flicker noise*. Crossovers from random telegraph noise and flicker noise were observed in point contacts between metals.

9.5 Continuous Markov processes. Fokker-Planck equation

Let us return to the problem of random walks and note that Eq. (9.3) can be reformulated for probabilities. We get

$$P(x,t|x_0,t_0) = \int dy\, P(x,t|y,\theta)P(y,\theta|x_0,t_0)\,, \quad t_0 < \theta < t\,. \tag{9.15}$$

This is the *Smoluchowski equation*. The instant θ can be arbitrary, so let us make it close to t. Namely, let us put $\theta = t - \tau$ and then tend τ to zero. We get

$$P(x,t|x_0,t_0) = \int dy\, P(x,t|y,t-\tau)P(y,t-\tau|x_0,t_0)\,. \tag{9.16}$$

Now let us multiply this equation by some arbitrary function $q(x)$ which has the properties

$$q(x) \to 0,\ q' \to 0 \quad \text{at} \quad x \to \pm\infty$$

and then integrate over x. We get

$$\begin{aligned}\int_{-\infty}^{\infty} dx\, q(x)P(x,t|x_0,t_0) &= \int_{-\infty}^{\infty} dy\, P(y,t-\tau|x_0,t_0)\int_{-\infty}^{\infty} dx\, q(x)P(x,t|y,t-\tau) \\ &= \int_{-\infty}^{\infty} dy\, P(y,t-\tau|x_0,t_0)\int_{-\infty}^{\infty} dx\, q(x)P(x,t|y,t-\tau) \\ &\times\left[q(y) + q'(y)(x-y) + q''(y)\frac{(x-y)^2}{2} = \cdots\right].\end{aligned}$$

Since

$$\int_{-\infty}^{\infty} dx\, P(x,t|y,t-\tau) = 1$$

we can replace in this integral $y \to x$ move it into l.h.s. Then we notice that

$$\frac{1}{\tau}\left[P(x,t|x_0,t_0) - P(x,t-\tau|x_0,t_0)\right] \approx \frac{\partial P(x,t|x_0,t_0)}{\partial \tau}.$$

As a result, we have

$$\int_{-\infty}^{\infty} dx\, q(x)\, \frac{\partial P(x,t|x_0,t_0)}{\partial \tau} = \int_{-\infty}^{\infty} dy\, P(y,t|x_0,t_0)\left[q'(y)A(y,t) + q''(y)\frac{B(y,t)}{2} + \cdots\right]. \tag{9.17}$$

Here

$$A(y,t) = \lim_{\tau=0} \frac{\langle x-y\rangle}{\tau} = \lim_{\tau=0} \frac{1}{\tau}\int_{-\infty}^{\infty} dx\,(x-y)P(x,t|y,t-\tau), \tag{9.18}$$

$$B(y,t) = \lim_{\tau=0} \frac{\langle (x-y)^2\rangle}{\tau} = \lim_{\tau=0} \frac{1}{\tau}\int_{-\infty}^{\infty} dx\,(x-y)^2 P(x,t|y,t-\tau). \tag{9.19}$$

Now we replace $y \to x$ in the r.h.s. of Eq. (9.17) and make integration by parts. Since $q(\pm\infty) = q'(\pm\infty) = 0$ and $q(x)$ is arbitrary we finally obtain

$$\frac{\partial P(x,t|x_0,t_0)}{\partial t} = -\frac{\partial}{\partial x}A(x,t)P(x,t|x_0,t_0) + \frac{1}{2}\frac{\partial^2}{\partial x^2}B(x,t)P(x,t|x_0,t_0). \tag{9.20}$$

This equation was derived in several contexts, so it has different names: *Einstein-Fokker equation, Fokker-Planck equation, 2nd Kolmogorov equation*. The diffusion equation (9.2) is a special case with constant A and B.

The solution of Eq. (9.20) must be non-negative and satisfy the initial condition

$$P(x,t_0|x_0,t_0) = \delta(x-x_0). \tag{9.21}$$

The physical interpretation of Eq. (9.20) is the following. Let a stream (ensemble) of particles emerge form the point x_0 at time t_0. Assume that they move independently. Then their concentration at the point x at time t is just $P(x,t|x_0,t_0)$. The the flow (mass current) S consists of the "convection" part AP, where A is the convection velocity, and of the diffusion part $-(1/2)(\partial BP/\partial x)$ where $B/2$ has a meaning of the diffusion coefficient,

$$S = AP - \frac{1}{2}\frac{\partial BP}{\partial x}. \tag{9.22}$$

We see that Eq. (9.20) is equivalent to the particle conservation,

$$\frac{\partial P}{\partial t} + \frac{\partial S}{\partial x} = 0.$$

Equation (9.22) is conventionally written as

$$S = \left(A - \frac{\partial B}{\partial x}\right)P(x,t|x_0,t_0) - \frac{B}{2}\frac{\partial P(x,t|x_0,t_0)}{\partial x}. \tag{9.23}$$

The combination $A - \partial B/\partial x$ is called the *drift velocity* while $B/2$ is the *diffusion constant*.

9.6 Fluctuations in a vacuum-tube generator

Consider a simplest generator of electric oscillations based on 3-electrode vacuum tube, see Fig. 9.5, the notations are clear from the picture.

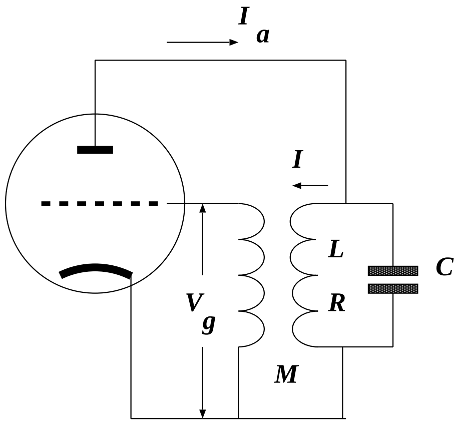

Figure 9.5: On the vacuum tube generator.

The tube acts as an amplifier, and the nonstationary part of the current is given by the expression

$$I_a = SV_g \left(1 - \frac{V_g^2}{2V_a^2}\right) \tag{9.24}$$

where V_0 is the DC anode voltage. The transformer facilitates feedback, such that

$$V_g = MI' \tag{9.25}$$

where M is the mutual inductance while prime means here the time derivative. Since the charge at the capacitor C is just

$$\tilde{Q} = \int^t dt'(I_a - I),$$

we have the following equation for the current through inductor L,

$$\begin{aligned} LI' + RI &= C^{-1} \int^t dt' \, [I_a(t') - I(t')] \quad \rightarrow \qquad (9.26) \\ I'' + \frac{R}{L} I'' + \frac{1}{LC} I &= \frac{SM}{LC} I' \left(1 - \frac{M^2 (I')^2}{3V_0^2}\right). \qquad (9.27) \end{aligned}$$

Now,it is convenient to introduce dimensionless variables,

$$\tau \equiv \omega_0 t, \; I_0 \equiv 2V_0/\omega_0 M, \; x \equiv I/I_0, \; \omega_0 MS \equiv \mu, \; \omega_0(MS - RC) \equiv p,$$

where $\omega_0 = (LC)^{-1}$ is the eigenfrequency of the circuit. Then Eq. (9.27) can be rewritten as

$$\ddot{x} + x = \mu \dot{x} \left(p - \frac{4}{x} \dot{x}^2\right). \tag{9.28}$$

Let us consider the situation when the feedback parameter μ is small and use the so-called method of slow perturbations. Namely, the solution of Eq. (9.28) at $\mu = 0$ is

$$x = r\,\cos(\tau+\varphi)\,, \tag{9.29}$$

and we will search the general solution in the form

$$x = r(\tau)\,\cos u(\tau)\,, \quad u(\tau) \equiv \tau + \varphi(\tau)\,, \tag{9.30}$$

assuming the both $r(\tau)$ and $\varphi(\tau)$ are "slow" functions. What does it mean will be clear a bit later. Then

$$\begin{aligned} \dot{x} &= \dot{r}\cos u - r\dot{u}\sin u\,, \\ \ddot{x} &= \frac{d}{d\tau}\left(\dot{r}\cos u - r\dot{u}\sin u\right) \approx -r\ddot{u}\sin u - 2\dot{r}\dot{u}\sin u - r\dot{u}^2\cos u\,. \end{aligned}$$

Here we neglected $\ddot{r}$ since $\mu \ll 1$. Now,

$$\dot{u} = 1+\dot{\varphi}, \quad \dot{u}^2 \approx 1 + 2\dot{\varphi}, \quad \ddot{u} \approx 0\,.$$

Then the l.h.s. of Eq. (9.28) is

$$-2\dot{r}\sin u - 2r\dot{\varphi}\cos u\,.$$

Since the r.h.s. of this equation is proportional to μ we have to substitute the derivatives up to the lowest order, as $\dot{x} = -r\sin u$ to get

$$\dot{x}\left(1-\frac{4}{3}\dot{x}^2\right) = -pr\sin u + \frac{4}{3}r^3\sin^3 u = -r(p-r^2)\sin u - \frac{r^2}{3}\sin 3u\,.$$

As a result we get,

$$-2\dot{r}\dot{u}\sin u - 2r\dot{\varphi}\cos u = -\mu r(p-r^2)\sin u - \mu\frac{r^2}{3}\sin 3u\,. \tag{9.31}$$

Now we can average this equation over "rapid motion". Namely, we multiply these equation by $\sin u(\tau)$ or by $\cos u(\tau)$ and then integrate over τ from 0 to 2π assuming $\varphi(\tau) = \text{const}$. From such a procedure we get

$$\dot{r} = \frac{\mu}{2}r(p-r^2)\,, \quad \dot{\varphi} = 0\,. \tag{9.32}$$

So up to our approximation, φ is constant. In Fig. 9.6 the function $r(p-r^2)$ is plotted for $p = -1, 0, 1$. In is clear that at $p \le 0$ there is only one stable point, $r = 0$. It can be shown that at $p > 0$ the only stable point $r = \sqrt{p}$.

Now we are prepared to write down the Fokker-Planck equation for conditional probability $P(r,\varphi,\tau|r_0,\varphi_0,\tau_0)$. In general, when we have several variables, the conditional probability can be expressed as a function of *vectors* $\mathbf{x} = \{x_\alpha\}$ and $\mathbf{x}_0 = \{x_{0\alpha}\}$. In a straightforward way we get vector $\mathbf{A}$ and tensor $\hat{B}$ as

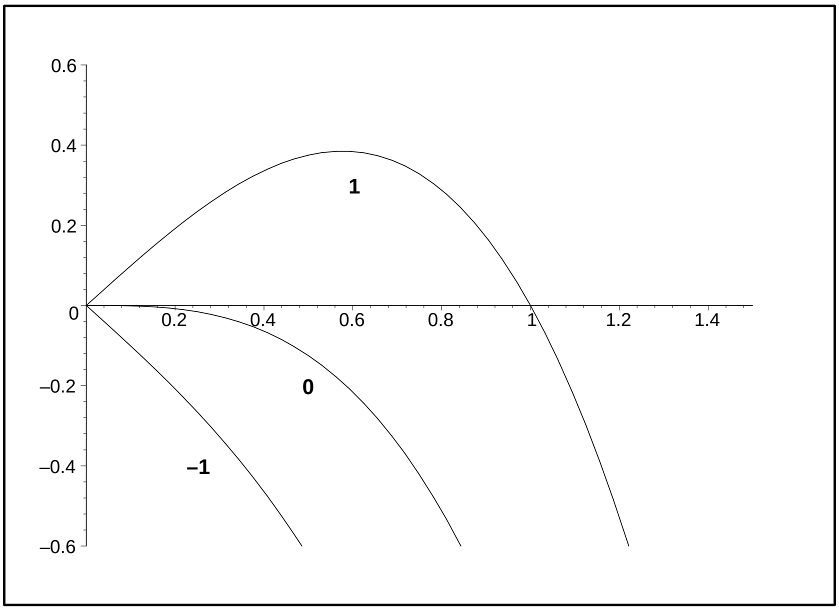

Figure 9.6: On the stationary points. The function $r(p-r^2)$ vs. r is plotted.

$$\begin{aligned}
\mathbf{A}(\mathbf{y},t) &= \lim_{\tau=0}\frac{\langle \mathbf{x}-\mathbf{y}\rangle}{\tau} = \lim_{\tau=0}\frac{1}{\tau}\int_{-\infty}^{\infty}(dx)\,(\mathbf{x}-\mathbf{y})P(\mathbf{x},t|y,t-\tau)\,, && (9.33)\\
\hat{B}(\mathbf{y},t) &= \lim_{\tau=0}\frac{\langle (x_i-y_i)(x_k-y_k)\rangle}{\tau}\\
&= \lim_{\tau=0}\frac{1}{\tau}\int_{-\infty}^{\infty}(dx)\,(x_i-y_i)(x_k-y_k)P(\mathbf{x},t|\mathbf{y},t-\tau)\,. && (9.34)
\end{aligned}$$

The Fokker-Planck equation acquires the form

$$\frac{\partial P(\mathbf{x},t|\mathbf{x}_0,t_0)}{\partial t} = -\mathrm{div}\,\mathbf{A}(\mathbf{x},t)P(\mathbf{x},t|\mathbf{x}_0,t_0) + \frac{1}{2}\sum_{ik}\frac{\partial^2}{\partial x_i \partial x_k}B_{ik}(\mathbf{x},t)P(\mathbf{x},t|\mathbf{x}_0,t_0)\,. \qquad (9.35)$$

Below we shall use a simple model of isotropic random forces with

$$B_{rr} = r^2 B_{\varphi\varphi} = \mu B \quad B_{r\varphi} = 0\,.$$

In our case,

$$A_r = \dot{r} = \frac{\mu}{2} r(p-r^2)\,, \quad A_\varphi = 0\,,$$

and we get

$$\mu^{-1}\frac{\partial P}{\partial t} = -\frac{\mu}{2}\frac{\partial[r(p-r^2)P]}{\partial r} + \frac{B}{2}\left[\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\frac{P}{r}\right) + \frac{1}{r^2}\frac{\partial^2 P}{\partial \varphi^2}\right]\,. \qquad (9.36)$$

It should be solved with the initial conditions

$$P(r,\varphi,t_0|r_0,\varphi_0,t_0) = \delta(r-r_0)\delta(\varphi-\varphi_0)\,.$$

Let us start from the stationary distribution

$$w(r,\varphi)=\lim_{t-t_0\to\infty}P(r,\varphi,t|r_0,\varphi_0,t_0)$$

for which the l.h.s. of previous equation is zero. If φ is chosen in the interval $(0,2\pi)$ all the values of φ are equal, and the solution is φ-independent. Then the equation (9.36) acquires the form

$$\frac{1}{2}\frac{\partial}{\partial r}\left(r(p-r^2)w-Br\frac{d}{dr}\frac{w}{r}\right)=0 \quad \text{or} \quad r(p-r^2)w-Br\frac{d}{dr}\frac{w}{r}=\text{const}\,. \tag{9.37}$$

The only solution which is regular at $r=0$ implies that the constant is equal to 0, and

$$w(r)\propto r\exp\left[\frac{1}{B}\int_0^r \rho d\rho\,(p-\rho^2)\right]=Cr\exp\left(-\frac{(p-r^2)^2}{4B}\right). \tag{9.38}$$

The proportionality coefficient is given by the normalization condition. Since w is φ-independent

$$C^{-1}=2\pi\int rdr\,\exp\left(-\frac{(p-r^2)^2}{4B}\right)=\pi^{3/2}\sqrt{B}\left[1+\text{erf}\left(\frac{p}{2\sqrt{B}}\right)\right].$$

Here

$$\text{erf}\,x=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt\,. \tag{9.39}$$

The profile $w(r)$ has a maximum at

$$r_m=\sqrt{\frac{1}{2}+\sqrt{\frac{p^2}{4}+B}}\approx\begin{cases}\sqrt{p}(1+B/2p^2) & \text{for} \quad p\gg 2\sqrt{B}\,,\\ \sqrt{B/|p|} & \text{for} \quad p\ll -2\sqrt{B}\,.\end{cases}$$

The function $w(r)$ is plotted in Fig. 9.7.

At negative x,

$$\text{erf}\,(-x)=-1+\frac{e^{-x^2}}{|x|\sqrt{\pi}}-\dots\,.$$

One can also neglect r^4 in the exponent of Eq. (9.38. Then we obtain

$$w(r,\varphi)\,dr\,d\varphi=\frac{|p|}{B}e^{-r^2|p|/2B}rdr\,\frac{d\varphi}{2\pi}\,,\quad \langle r^2\rangle=2B|p| \tag{9.40}$$

which is called the *Rayleigh distribution*.

If the generator is strongly excited, $p\gg 2\sqrt{B}$, we can put $r=\sqrt{p}+\rho$ and neglect ρ^2 in Eq. (9.38). Then we obtain,

$$w(r,\varphi)\,dr\,d\varphi=\sqrt{\frac{p}{\pi B}}e^{-\rho^2p/B}d\rho\,\frac{d\varphi}{2\pi}\,,\quad \langle\rho^2\rangle=B/2p\,, \tag{9.41}$$

i.e. the Gaussian distribution.

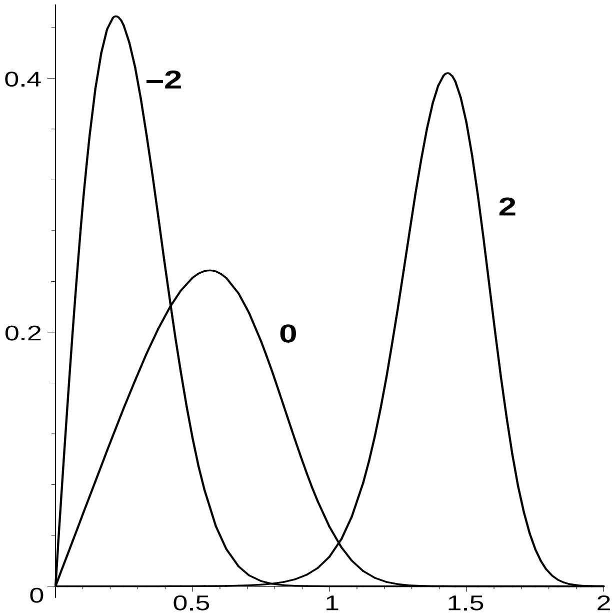

Figure 9.7: Plots of the distribution $w(r)$ for different values of $p/\sqrt{B}$, shown near the curves.

To find the generation spectrum let us consider the case of large amplitudes,

$$r_m = \sqrt{p} \gg \sqrt{\langle \rho^2 \rangle} = \sqrt{B/2p}\,.$$

In this case we can linearize the dynamic equations (9.32) with respect to $\rho = r - \sqrt{p}$ to get

$$\dot{\rho} = -p\rho\,, \quad \dot{\varphi} = 0\,. \tag{9.42}$$

The simplified Fokker-Planck equation acquires the form

$$\mu^{-1}\frac{\partial P}{\partial t} = p\frac{\partial \rho P}{\partial \rho} + \frac{B}{2}\left(\frac{\partial^2 P}{\partial \rho^2} + \frac{1}{p}\frac{\partial^2 P}{\partial \varphi^2}\right)\,. \tag{9.43}$$

This equation we can analyze for arbitrary initial conditions. Denoting $P(\rho,\phi,t|\rho_0,\phi_0,t_0) \equiv P(\rho,\phi|\rho_0,\phi_0,t-t_0)$ we immediately get that fluctuations of ρ and φ are statistically-independent,

$$P(\rho,\phi|\rho_0,\phi_0,t-t_0) = P_\rho(\rho|\rho_0,t-t_0)\,P_\varphi(\varphi|\varphi_0,t-t_0)\,. \tag{9.44}$$

The fluctuations in the amplitude are Gaussian,

$$P_\rho(\rho|\rho_0,t-t_0) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(\rho-\bar{\rho})^2/2\sigma^2} \tag{9.45}$$

with

$$\bar{\rho} = \rho_0\,e^{-\mu p(t-t_0)}\,, \quad \sigma^2 = \frac{B}{2p}\left(1 - e^{-2\mu p(t-t_0)}\right)\,. \tag{9.46}$$

The $\bar{\rho}(t)$ is just the dependence due to mechanical equations of motion (9.32) linearized near $r = \sqrt{p}$. The solution for $P_\varphi(\varphi|\varphi_0,t-t_0)$ has the form

$$P_\varphi(\varphi|\varphi_0,t-t_0) = \frac{1}{2\pi}\left[1 + 2\sum_{n=1}^{\infty} e^{-n^2(B/2p)\mu(t-t_0)}\cos n(\varphi - \varphi_0)\right]\,. \tag{9.47}$$

As $t-t_0$ increases, the average deviation from the stationary solution, $\bar{\rho}$, decays while dispersion σ grows. One can easily check that the $t-t_0$ the system tends to stationary distributions. At $t=t_0+0$ it gives proper initial conditions since

$$1+\sum_{n=1}^{\infty}\cos n(\varphi-\varphi_0)=2\pi\delta(\varphi-\varphi_0)\,.$$

Having the conditional probabilities (9.46) and (9.47) one can find all necessary correlation functions. Indeed, the *joint probability*, $w(\rho_\tau,\rho,\tau$, to find the values ρ and ρ_τ after time delay τ is

$$\begin{aligned}
w(\rho_\tau,\rho,\tau) &= w(\rho)P(\rho_\tau|\rho,\tau)=\frac{1}{2\pi\sigma_\infty\sigma_\tau}\exp\left(-\frac{\rho^2}{2\sigma_\infty^2}-\frac{(\rho_\tau-\bar{\rho}_\tau)^2}{2\sigma_\tau^2}\right)\,, && (9.48)\\
w(\varphi_\tau,\varphi,\tau) &= w(\varphi)P(\varphi_\tau|\varphi,\tau)=\frac{1}{4\pi^2}\left[1+2\sum_{n-1}^{\infty}e^{-n^2D\tau}\cos n(\varphi_\tau-\varphi)\right]\,. && (9.49)
\end{aligned}$$

Here

$$\bar{\rho}_\tau=\rho\, e^{-p\mu\tau}\,,\quad \sigma_\tau^2=\frac{B}{2p}\left(1-e^{-2p\mu\tau}\right)\,,\quad D=\mu B/2p\,.$$

using these equations let us calculate the correlation function

$$\psi_\tau\equiv\langle x(0)x(\tau)\rangle\,. \qquad (9.50)$$

Lets us for brevity denote $x(0)=x$, $x(\tau)=x_\tau$. Since

$$x(t)=[\sqrt{p}+\rho(t)]\cos[t+\varphi(t)]$$

and $\langle x(t)\rangle=0$ due to the facts that $\langle\rho\rangle$=0, ρ and φ are statistically independent at any time, and the phase distribution is uniform. Hence $\langle\cos\varphi\rangle=\langle\sin\varphi\rangle=0$. As a result,

$$\begin{aligned}
\psi_\tau &= \langle(\sqrt{p}+\rho)(\sqrt{p}+\rho_\tau)\rangle\langle\cos(t+\varphi)\cos(t+\tau+\varphi_\tau)\rangle\\
&= \frac{1}{2}\left(p+\langle\rho\rho_\tau\rangle\right)\left[\langle\cos(\tau+\varphi_\tau-\varphi)\rangle+\langle\cos(2t+\tau+\varphi_\tau+\varphi)\rangle\right]\,. && (9.51)
\end{aligned}$$

Then, from Eq. (9.48) we get,

$$\langle\rho\rho_\tau\rangle=\int d\rho_\tau d\rho\,\rho_\tau\rho\, w(\rho_\tau,\rho,\tau)=De^{-p\mu|\tau|}\,. \qquad (9.52)$$

Since the joint probability $w(\varphi_\tau,\varphi,\tau)$, Eq. (9.49) is an even function of $\varphi_\tau-\varphi$ one can easily show that

$$\langle\cos(\varphi_\tau+\varphi)\rangle=\langle\sin(\varphi_\tau\pm\varphi)\rangle=0\,.$$

As a result, $\langle\cos(\tau+\varphi_\tau-\varphi)\rangle=\langle\cos(\varphi_\tau-\varphi)\rangle\cos\tau$. Then,

$$\begin{aligned}
\langle\cos(\varphi_\tau-\varphi)\rangle &= \frac{1}{4\pi^2}\int_0^{2\pi}d\varphi_\tau\int_0^{2\pi}d\varphi\,\cos(\varphi_\tau-\varphi)\left[1+2\sum_{n=1}^{\infty}e^{-n^2D\mu|\tau|}\cos n(\varphi_\tau-\varphi)\right]\\
&= e^{-D\mu|\tau|}\,. && (9.53)
\end{aligned}$$

Collecting all the factors we finally obtain:

$$\psi_\tau = \frac{1}{2}\left(p + De^{-p\mu|\tau|}\right) e^{-D\mu|\tau|}\cos\tau\,. \tag{9.54}$$

The first factor in given by the fluctuations in the amplitude while the second is due to the phase fluctuations. The oscillation spectrum is given by Fourier transform of the above function. In the dimensional variables, $\tau \to \omega_0\tau$, we get

$$\psi_\omega = \Phi(\omega-\omega_0) + \Phi(\omega+\omega_0) \approx \Phi(\omega-\omega_0)$$

with

$$\Phi(\alpha) = \frac{p}{2\pi}\left[\frac{\mathcal{D}}{\alpha^2+\mathcal{D}^2} + \frac{D}{h}\cdot\frac{\mathcal{D}+h}{\alpha^2+(\mathcal{D}+h)^2}\right]. \tag{9.55}$$

Here

$$\mathcal{D} = \mu D\omega_0, \quad h = \mu p\omega_0$$

are dimensional phase diffusion constant and increment. That means a superposition of a narrow line with the half-width $\mathcal{D}$ with a broad line with the half-width $\mathcal{D}+h$.

Chapter 10

Non-Ideal Gases

10.1 Deviation of gases from the ideal state

We start from the case of a weakly non-ideal gas where the only pair interaction is important. The simplest case is the monoatomic gas where the interaction potential depends only on the distance between the atoms. Since

$$\mathcal{H}(p,q)=\sum_a \frac{p_a^2}{2m}+V$$

we can write the free energy as

$$\begin{aligned} F &= F_{\text{id}}+F_i\,, \\ F_i &= -T\,\log\left[\frac{1}{V^N}\int\cdots\int e^{-\beta V(q)}\prod_a dV_a\right]. \end{aligned}$$

Here F_{id} is the free energy of the ideal gas. The factor V^{-N} in the argument of the logarithm is just because for the ideal gas the integral over the volume is V^N. It is convenient to rewrite the above expression in the form

$$F=F_{\text{id}}-T\,\log\left[1+\frac{1}{V^N}\int\cdots\int\left(e^{-\beta V(q)}-1\right)dV_1\ldots dV_N\right].$$

Now let us make use of the assumption that the gas is rarefied and only two atoms can collide at the same time. Then the integrand is not small only if some two atoms are close together. Such a pair can be extracted in $N(N-1)/2$ ways. Thus the integral can be expresses as

$$\frac{N(N-1)}{2V^N}\int\cdots\int\left(e^{-\beta V_{12}}-1\right)dV_1\ldots dV_N\approx\frac{N^2}{2V^2}\int\left(e^{-\beta V_{12}}-1\right)dV_1\,dV_2\,.$$

Here V_{12} is the pair interaction potential. Since $\log(1+x)\approx x$ at $x\ll 1$ and the gas density N/V is assumed to be small,

$$F_i=-\frac{TN^2}{2V^2}\int\left(e^{-\beta V_{12}}-1\right)dV_1\,dV_2\,.$$

Since the interaction depends only on the *distance between the atoms* we can integrate out the center of mass and get V. In this way we are left with the expression

$$F_i = \frac{TN^2}{V} B(T), \quad B(T) = \frac{1}{2} \int \left(1 - e^{-\beta V_{12}}\right) dV,$$

where dV stands for relative coordinates. Since $P = -\partial F/\partial V$ we obtain the following correction to the equation of state,

$$P = \frac{NT}{V}\left[1 + \frac{N}{V}B(T)\right].$$

The corrections to other thermodynamic potentials can be found using the principle of small increments. For example, $G_i = NBP$.

A typical dependence of the interaction potential energy is shown in Fig. 10.1

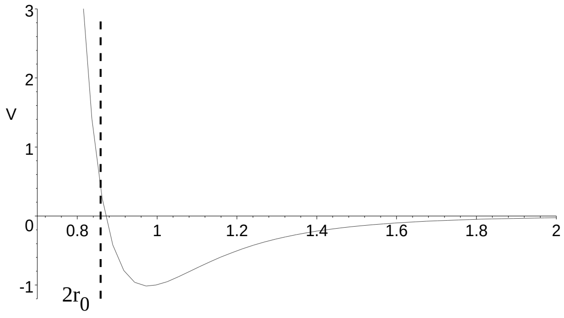

Figure 10.1: A typical dependence of the normalized potential energy V/V_0 versus the normalized interatomic distance r/r_0.

Expansion in powers of density

In principle, the equation of state can be expressed in the form of expansion in the density,

$$P = T \sum_{n=1}^{\infty} \left(\frac{N}{V}\right)^n B_n(T), \quad B_1 = 1$$

where coefficients $B_n(T)$ are called the *virial coefficients*. To determine them it is convenient to use the Landau fee energy, $\Omega = -PV$,

$$e^{\beta\Omega} = \sum_{N=0}^{\infty} \frac{1}{N!} e^{\beta\mu N} \int d\Gamma_N e^{-\beta\mathcal{H}_N(p,q)}.$$

For the N-particle Hamiltonians we have,

$$\mathcal{H}_1 = \frac{p^2}{2m}, \quad \mathcal{H}_2 = \sum_{a=1,2} \frac{p_a^2}{2m} + V_{12}, \quad \mathcal{H}_3 = \sum_{a=1}^{3} \frac{p_a^2}{2m} + V_{123}, \quad \ldots$$

where in general

$$V_{123} \neq V_{12} + V_{13} + V_{23}.$$

If one denotes

$$\zeta \equiv \frac{e^{\beta\mu}}{(2\pi\hbar)^3}\int d^3p\, e^{-\beta p^2/2m} = \left(\frac{mT}{2\pi\hbar^2}\right)^{3/2} e^{\beta\mu}$$

the expression for Ω can be written as

$$\Omega = -PV = -T\log\left[1+\zeta V + \frac{\zeta^2}{2!}\int e^{-\beta V_{12}}\, dV_1 dV_2 + \frac{\zeta^3}{3!}\int e^{-\beta V_{123}}\, dV_1 dV_2 dV_3 + \dots\right].$$

Again, we can integrate out the center-of-mass motion to get

$$\Omega = -PV = -T\log\left[1+\zeta V + \frac{\zeta^2 V}{2!}\int e^{-\beta V_{12}}\, dV_2 + \frac{\zeta^3 V}{3!}\int e^{-\beta V_{123}}\, dV_2 dV_3 + \dots\right].$$

We arrive at the expansion in powers of ζ

$$P = T\sum_{n=1}^{\infty}\frac{J_n}{n!}\zeta^n$$

with

$$\begin{aligned} J_1 &= 1, \quad J_2 = \int\left(e^{-\beta V_{12}}-1\right) dV_2 \\ J_3 &= \int\left(e^{-\beta V_{123}} - e^{-\beta V_{12}} - e^{-\beta V_{23}} - e^{-\beta V_{13}} + 2\right) dV_2 dV_3, \quad \dots \end{aligned}$$

The structure of the integrals J_n is clear, they are not small only if n particles are close together. To obtain equation of state one has to relate the parameter ζ to the density. Since

$$N = -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = V\left(\frac{\partial P}{\partial\mu}\right)_{T,V}$$

we have

$$\frac{N}{V} = \sum_{n=1}^{\infty}\frac{J_n}{(n-1)!}\zeta^n .$$

The set of equations $P(\zeta)$ and $N(\zeta)$ defines the equation of state.

Spatial correlation of density fluctuations

Let us denote $n(\mathbf{r})$ the fluctuating number density, $n\,dV$ being the particle number in the element dV. The correlation in the positions of the particles can be characterized by the average

$$\langle \Delta n(\mathbf{r}_1)\,\Delta n(\mathbf{r}_2)\rangle = \overline{n(\mathbf{r}_1)n(\mathbf{r}_2)} - \bar{n}^2 .$$

For a uniform system it can depend only on the difference $r = |\mathbf{r}_1 - \mathbf{r}_2|$, it must decay at $r \to \infty$.

The product $\langle n(\mathbf{r}_1)\, n(\mathbf{r}_2)\rangle$ can be expressed through the probability density w_{12} to find the particle 2 at the distance r from the particle one. Since the total probability of such event is $\bar{n}\, w_{12}(r)\, dV_2$ we might write

$$\langle n(\mathbf{r}_1)\, n(\mathbf{r}_2)\rangle = \bar{n}^2 w_{12}\,.$$

That would be a mistake because it ignores the fact that if the points 1 and 2 coincide, the particle belongs both to dV_1 and dV_2. Thus the correct formula is

$$G(\mathbf{r}_1-\mathbf{r}_2) = \langle n(\mathbf{r}_1)\, n(\mathbf{r}_2)\rangle = \bar{n}^2 w_{12} + \bar{n}\,\delta(\mathbf{r}_1-\mathbf{r}_2)\,.$$

In a similar way,

$$\langle \Delta n(\mathbf{r}_1)\, \Delta n(\mathbf{r}_2)\rangle = \bar{n}\nu(r) + \bar{n}\,\delta(\mathbf{r}_1-\mathbf{r}_2)\,, \quad \nu(r) \equiv \bar{n}[w_{12}(r)-1]\,.$$

The quantities $\langle \Delta n(\mathbf{r}_1)\, \Delta n(\mathbf{r}_2)\rangle$ and $\nu(r)$ are called the *correlation functions*. Integrating the previous equation with respect to $dV_1\, dV_2$ over a volume V we have to recover $\langle (\Delta N)^2\rangle$. From the formula above we find the normalization condition,

$$\int \nu\, dV = \frac{\langle (\Delta N)^2\rangle}{N} - 1 = -\frac{TN}{V^2}\left(\frac{\partial V}{\partial P}\right)_T - 1\,.$$

This is zero for an ideal gas, only interaction leading to a finite compressibility leads to a correlation. Since for the pair interaction

$$\nu(r) = \bar{n}\left(e^{-\beta V_{12}(r)} - 1\right)$$

we get

$$J_2 = \int \left(e^{-\beta V_{12}(r)} - 1\right) dV = \frac{V}{N}\int \nu(r)\, d^3 r\,.$$

We see that the correlation function is an important property.

It can be directly determined from the experiments on the light scattering. Consider a wave

$$\mathbf{F}(\mathbf{R}) = \mathbf{F}_0\, e^{i\mathbf{k}\mathbf{R}-i\omega t}$$

which is scattered by the particles residing at points $\mathbf{r}_i$ (see Fig. 10.2). Let the scattering amplitude f be the same for all the particles. Then the scattered wave at the detection point $\mathbf{r}$ is

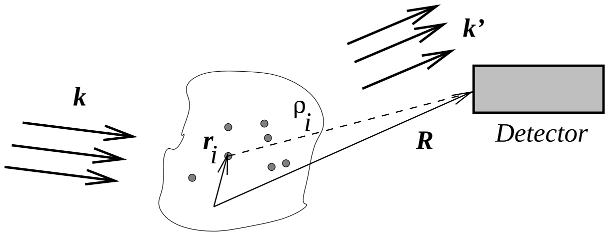

Figure 10.2: On the scattering problem.

$$\mathbf{F}_{sc}(\mathbf{r}) = f\sum_{i=1}^{N} \mathbf{F}(\mathbf{r}_i)\frac{e^{ik\rho_i}}{\rho_i} = f\mathbf{F}_0 \sum_{i=1}^{N} e^{i(\mathbf{k}\mathbf{r}_i-\omega t)}\,\frac{e^{ik\rho_i}}{\rho_i}\,.$$

Here $\rho_i = |\mathbf{R} - \mathbf{r}_i|$ is the distance between the scatterer and the detector. We have $\rho_i = R - r_i \cos\phi$ where ϕ is the angle between $\mathbf{r}_i$ and $\mathbf{R}$ Now let us take into account the fact that the detector is usually place far from the scattering medium, $R \gg r_i$. Under this approximation we can replace ρ_i by R in the denominator of the previous formula, while the phase needs more exact treatment,

$$\mathbf{k}\mathbf{r}_i + k\rho_i = kr_i + kR - kr_i \cos\phi = kR - \mathbf{r}_i(\mathbf{k} - \mathbf{k}') .$$

In this way we get,

$$\mathbf{F}_{sc} = f\mathbf{F}_0 \frac{e^{ikR - i\omega t}}{R} \sum_i e^{-i\mathbf{r}_i\mathbf{q}} , \quad \mathbf{q} \equiv \mathbf{k} - \mathbf{k}' .$$

Thus the ratio between the average intensity of the scattered wave and the intensity of the incident wave is

$$|f|^2 \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\mathbf{q}(\mathbf{r}_i - \mathbf{r}_j)} \equiv N|f|^2 S(\mathbf{q}) .$$

$S(\mathbf{q})$ is called *the structure factor*. Replacing the double sum by its ensemble average, we can write it as

$$S(\mathbf{q}) = \frac{1}{N} \int dV_1 \int dV_2 \, \langle n(\mathbf{r}_1) n(\mathbf{r}_2) \rangle e^{i\mathbf{q}(\mathbf{r}_1 - \mathbf{r}_2)} = 1 + \int d^3r \, \nu(r) \, e^{i\mathbf{q}\mathbf{r}} .$$

This is the way to extract the correlation function from the experiment on light scattering. Note that for small scattering angles $q \to 0$ the structure factor is proportional to the compressibility of the system. Thus one can expect singularities of the scattering cross section at the phase transition point.

Van der Waals' formula

There is an interpolation formula which allows to follow the phase transition between gaseous and liquid state. To derive such a formula let us assume that the interaction is small in a sense that $V_0 \ll T$. To perform the integration for the virial coefficient let us assume that at $r \lesssim 2r_0$ we have $V_{12} > T$.The quantity $2r_0$ has a meaning of the atomic diameter. Let us split the integral for $B(T)$ into 2 parts,

$$B(T) = 2\pi \int_0^{2r_0} \left(1 - e^{-\beta V_{12}}\right) r^2 \, dr + 2\pi \int_{2r_0}^{\infty} \left(1 - e^{-\beta V_{12}}\right) r^2 \, dr .$$

In the first integral we neglect the exponential and obtain for that $b = 16\pi r_0^3/3$. In the second integral we expand the exponential to get

$$-2\pi\beta \int_{2r_0}^{\infty} |V_{12}| \, r^2 \, dr \equiv -2\beta a .$$

In this way,

$$B(T) = b - a/T \quad \to \quad F_i = N^2(bT - a)/V \, , \; G_i = NP(b - a/T) .$$

The total free energy is

$$F = Nf(T) - NT\,\log(e/N) - NT(\log V - Nb/V) - N^2a/V\,.$$

Now we do something approximate. Since we have assumed that the gas is rarefied,

$$V \gg Nb \quad \rightarrow \quad \log V - Nb/V \approx \log(V - Nb)\,.$$

Let us assume that this expression holds also for large density! As a result,

$$F_i = -NT\,\log(1 - Nb/V) - N^2a/V\,.$$

This formula take account of the *finite compressibility* at large density. Now we get

$$P = -\frac{\partial F}{\partial V} = \frac{NT}{V - Nb} - \frac{N^2a}{V^2} \quad \rightarrow \quad \left(P + \frac{N^2a}{V^2}\right)(V - Nb) = NT\,.$$

This the *van der Waals' equation*.

Chapter 11

Phase Equilibrium

11.1 Conditions for phase equilibrium

Consider an isolated body with the energy E and volume V. This state is non necessary *homogeneous*, it can consist of two (or more) phase. For the case of two phases, the equilibrium conditions are

$$T_1 = T_2\,, \quad P_1 = P_2\,, \quad \mu_1 = \mu_2\,.$$

Since the temperature and pressure remain constant, the equality

$$\mu_1(P,T) = \mu_2(P,T)$$

correspond to a line in the (P,T)-plane along which the phases coyexist. In the (T,V)-plane, there are 2 lines because at a given pressure and temperature can have different volumes. This is illustrated in Fig. 11.1

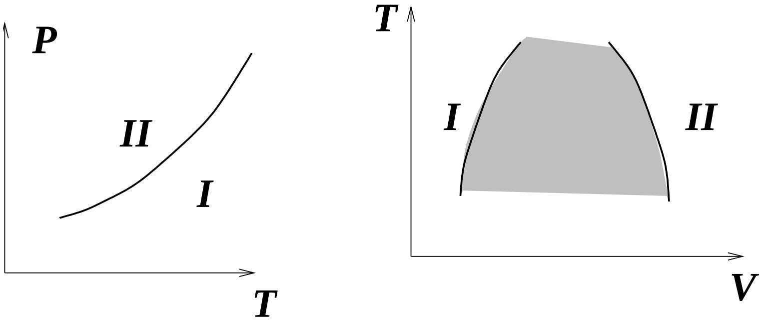

Figure 11.1: The hatched area corresponds to a mixture of 2 phases.

The Clapeyron-Clausius formula

Differentiating the equilibrium condition $\mu_1 = \mu_2$ with respect to temperature we get

$$\frac{\partial \mu_1}{\partial T} + \frac{\partial \mu_1}{\partial P}\frac{dP}{dT} = \frac{\partial \mu_2}{\partial T} + \frac{\partial \mu_2}{\partial P}\frac{dP}{dT}\,.$$

Since

$$\left(\frac{\partial \mu}{\partial T}\right)_P = -s\,, \quad \left(\frac{\partial \mu}{\partial P}\right)_T = v \quad (s \equiv S/N, v \equiv V/N)$$

we obtain

$$\frac{dP}{dT} = \frac{s_1 - s_2}{v_1 - v_2} = \frac{q}{T(v_1 - v_2)}$$

where $q = T(s_1 - s_2)$ is the heat of transition. This the Clapeyron-Clausius formula. This equation gives the temperature dependence of pressure along the phase equilibrium curve. Written as

$$\frac{dT}{dP} = \frac{T(v_1 - v_2)}{q}$$

it gives the change in the equilibrium temperature (freezing point, or boiling point) when the pressure changes.

The law of corresponding states

Let us discuss as an example the van der Waals equation of state,

$$P = \frac{NT}{V - Nb} - \frac{N^2 a}{V^2}\,.$$

We can notice that there is a special temperature, T_c, at which the isotherm has an *inflection point*. We have:

$$\left(\frac{\partial P}{\partial V}\right)_{T_c} = 0, \; \left(\frac{\partial^2 P}{\partial V^2}\right)_{T_c} = 0 \quad \rightarrow \quad T_c = \frac{8}{27}\frac{a}{b}, \; V_c = 3Nb, \; P_c = \frac{1}{27}\frac{a}{b^2}\,.$$

Measuring the P, V, and T in the units P_c, V_c and T_c, respectively, we can rewrite the equation of state in the form

$$\left(P' + \frac{3}{V'^2}\right)(3V' - 1) = 8T'\,.$$

The van der Waals' isotherms determined by this equation are shown in Fig. 11.2. So we are able to express the equation of state in a form which is independent of the particular properties of the gas. We observe that at $T < T_c$ there are regions with positive $\partial P/\partial V$. In that regions the system cannot be stable. On the other hand, we know that the equilibrium can be achieved only at $P = \text{const}$. As a result we arrive at the situation shown in Fig. 11.3 To the right from the point e the system is in a gas state, while to the left from the pint b it is in a liquid state. The phase equilibrium takes place along the straight line $b - e$ which corresponds to $P = \text{const}$. This line must be constructed in a way to keep $\mu_1 = \mu_2$. This difference can be expressed as an integral along the transition path,

$$\mu_e - \mu_b = \int_b^e d\mu = 0\,.$$

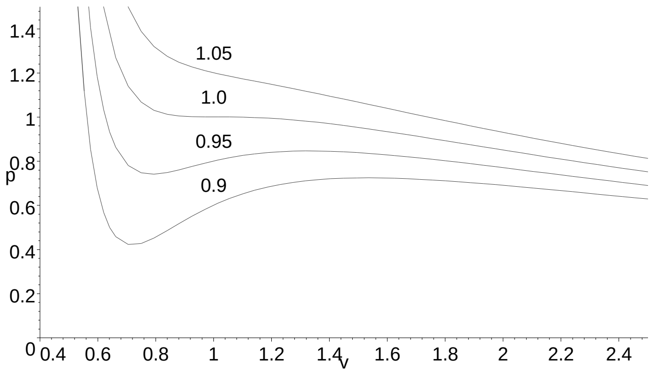

Figure 11.2: Van der Waals' isotherms. The numbers near the curves are the ratios T/T_c.

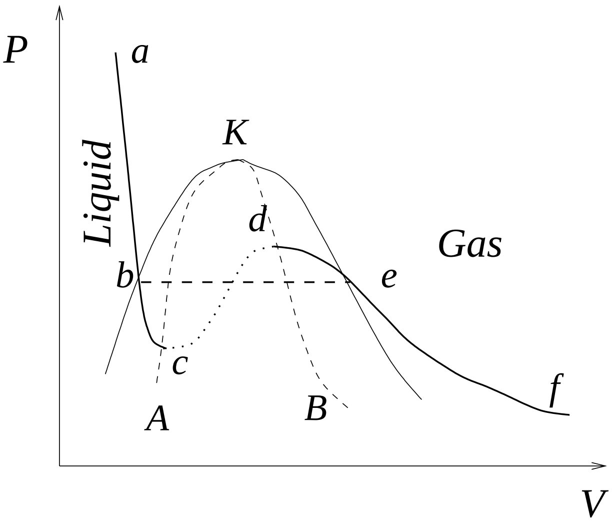

Figure 11.3: On the gas-liquid transition.

Since $d\mu = (V/N)\,dP$ we come to the conclusion that the integral along the isotherm,

$$\int_b^e V\,dP = 0\,.$$

Geometrically it means that the areas between the isotherm and the line $b-e$ below and above the line must be *equal*. This rule is called the *Maxwell construction*.

Van der Waals theory of the critical point

The calculations can be easily performed near the critical point, T_c. Introducing notations

$$P = P_c(1+p), \quad T = T_c(1+\tau) \quad \frac{N}{V} = \frac{N}{V_c}(1+n)$$

and expanding in powers of τ and n up to non-trivial order we arrive at the equation of state

$$p = 4\tau + 6\tau n + (3/2)n^3\,, \quad (\partial p/\partial n)_\tau = 6\tau + (9/2)n^2\,.$$

The Maxwell condition can be written as

$$\int_b^e V\,dP = 0 \quad \rightarrow \quad \int_{n_b}^{n_e} n \left(\frac{\partial p}{\partial n}\right)_\tau dn = 0\,.$$

Since the integrand is an odd function of n we get $n_e = -n_b$. From the condition $p_1 = p_2$ we obtain

$$n_e = -n_b = 2\sqrt{-\tau}\,, \quad p = 4\tau\,.$$

The boundary of the metastable region are given by the equation

$$n_d = -n_c = 2\sqrt{-\tau/3}\,.$$

According to the Clapeyron-Clausius formula, the heat of transition is proportional to $\sqrt{-\tau}$.

Chapter 12

Continuous Phase Transitions

Introduction

In the following set of lectures we shall discuss so-called continuous phase transitions, or transitions of the 2nd order. A generic system to discuss the transitions is a magnet. Thus we start in recollection of thermodynamics of a magnet.

Thermodynamics of a magnet

Consider a long and thin solenoid with the length L and number of windings N. Then the magnetic field inside the solenoid is $H = NI/L$, where I is the electric current. When the magnetic flux Φ

through the solenoid is changing an *electro-motive force*, $\mathcal{E} = -\partial\Phi/\partial t$, the work made by the battery being

$$\Delta W_b = \mathcal{E} I \, \Delta t = I \, \Delta\Phi = NAI \, \Delta B = VH\Delta B \, .$$

Here A is the solenoid area, V is its volume, while $B \equiv \Phi/A$ is the magnetic induction. Introducing *magnetization* $\tilde{M}$ from the relationship $B = \mu_0(H + \tilde{M})$ we can express the work per volume as

$$\Delta W_b/V = \Delta(\mu_0 H^2/2) + \mu_0 H \, \Delta\tilde{M} \, .$$

The first item is the energy of external magnetic field which is present also in the absence of material. The second term is the work done by the field *on* the material. Consequently, the work done *by the material* can be expressed through the total magnetic moment $M \equiv V\tilde{M}$ as

$$\Delta W = -H \, \Delta M \, .$$

As a result, the 1st law of thermodynamics reads as

$$dU = T \, dS - H \, dM$$

and

$$T = \left(\frac{\partial U}{\partial S}\right)_M , \quad H = \left(\frac{\partial U}{\partial M}\right)_S .$$

Generalizing of other thermodynamic potentials is more or less obvious:

$$\begin{aligned} F(T,M) &= U - TS \rightarrow dF = -S\,dT + H\,dM\,,\ S = -\left(\frac{\partial F}{\partial T}\right)_M,\ H = \left(\frac{\partial F}{\partial M}\right)_T; \\ G(T,H) &= F - HM \rightarrow dF = -S\,dT - M\,dH\,,\ S = -\left(\frac{\partial G}{\partial T}\right)_H,\ M = -\left(\frac{\partial G}{\partial H}\right)_T. \end{aligned}$$

General considerations

Uniaxial ferromagnet is a simplest example of phase transition. At zero external magnetic field magnetization vanishes at certain temperature, T_c. Below T_c a *spontaneous* magnetization $M_0(T)$ is observed, its direction can be "up" and "down" with equal likelihood. The dependence $M_0(T)$ is schematically shown in Fig. 12.1. This figure can be explained in the framework the Ising model. According to that model, spins are located at lattice points in a regular lattice. They can point up and down. Nearest neighbors interact in a way that is is favorable to point in the same direction. As we know, the free energy $G = U - TS - HM$ has to be minimized. At $H = 0$ it consists of two terms, U and $-T\,S(T)$. At low temperatures the energy U is more important and the system is a ferromagnet. At high temperature the free energy is minimized by *disordered* state in which the entropy is large. It is the fight between the order and disorder that makes the critical state specifically interesting.

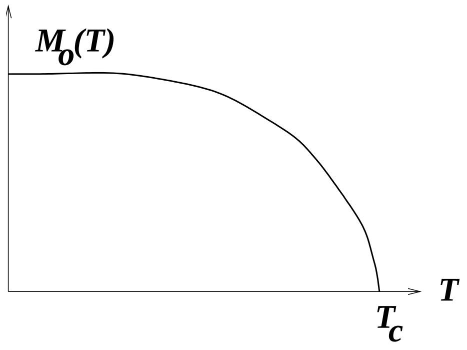

Figure 12.1: The spontaneous magnetization as a function of temperature.

In real life, the above conclusion is true for the dimension $d \geq 2$, while for $d = 1$ the magnetization never vanishes. That indicates that phase transition is a collective effect depending on interaction of many spins. The higher the dimension the greater the number of neighbors, the stronger the collective effects.

Coming back to the case of a uniaxial ferromagnet, we can define the spontaneous magnetization $M_0(T)$ as the *order parameter*. In some cases it is difficult to specify the order parameter. In general it can be a scalar, a vector, a tensor, etc.

When an order parameter is specified, the next question is if the transition is of 1st order, or it is *continuous*. In the example of the ferromagnet, it is the question if the spontaneous magnetization goes to zero at $T \rightarrow T_c^-$ is a continuous way, or it is abrupt. This is a delicate issue. Assume that we ramp the external magnetic field from a negative to a positive value. Than the $H = 0$ the magnetic moment changes from $-M_0(T)$ to $M_0(T)$, i. e. the transition is of the 1st order. Thus the character of transition cannot be specified without specifying the *transition path*.

At continuous phase transitions *critical phenomena* are observed. Usually a number of quantities show *power law* behavior close to T_c, e. g.

$$M_0(T) \sim (T_c - T)^{\beta}\,.$$

Other examples of critical exponents for thermodynamic quantities are

$$C \sim |T_c - T|^{-\alpha}, \quad \chi = \left(\frac{\partial M}{\partial H}\right)_T \sim |T_c - T|^{-\gamma}, \quad |M(H, T = T_c)| \sim H^{1/\delta}\,.$$

An important concept of the theory of continuous phase transitions is the *universality class*. It became clear during 1960's that a number of details in a given system are irrelevant as far as critical exponents are concerned. Systems of the same dimension and with a phase transition into an ordered state with the same symmetry belong to the same *universality class*. They have essentially the same critical properties. The justification of that concept will be discussed later.

Examples of phase transitions

We shall classify the phase transitions by the spatial dimensionality and the number of independent components of the order parameter, (d, n).

Magnetic systems

Uniaxial ferromagnets: described by the Ising model, the universality class being $(d, 1)$.

Hard axis ferromagnets: $\mathbf{m}_0$ is a vector in the *xy*-plane perpendicular to the hard axis, $(d, 2)$. This model is called the *isotropic xy*-model, the universality class being $(d, 2)$.

Heisenberg model: the magnetization is isotropic in 3 dimensions, the universality class being $(d, 3)$.

So far we have assumed isotropy in spin space. This assumption can be relaxed, and one can also expect, say, cubic anisotropy. In this case the corners of a (hyper)cube in spin space are energetically favored (or suppressed). That gives rise to *new* universality classes.

Antiferromagnets: Neighboring spins prefer to point to different directions, and on the average there are equally many spins pointing in opposite directions. The order parameter on a 3d cubic lattice can be chosen as *staggered* magnetization,

$$\mathbf{m}_s = \sum_{klm} (-1)^{k+l+m} \langle \mathbf{m}_{klm} \rangle\,,$$

where $\langle \mathbf{m}_{klm} \rangle$ is the magnetic moment of spin at the lattice point (k, l, m). Note that the universality classes for antiferromagnets are the same as for ferromagnets.

Fluid systems

The universality class is the same as for the Ising model, $(d,1)$ because the order parameter is a scalar. The situation is more complex for liquid mixtures.

Superfluids and Superconductors

At the temperature $T = T_\lambda \approx 2K$, liquid He^4 undergoes a continuous phase transition from the normal phase (HeI) to a (partially) superfluid phase (HeII). The relative fraction of superfluid component, ρ_s/ρ decreases continuously from 1 at $T = 0$ to 0 at $T = T_\lambda$. It appears that the proper order parameter is a *two-component* rotation invariant quantity, so the universality class in $(3,2)$.

A similar situation is relevant to superconductors. However, the critical region in the ”old” superconductors is so narrow that it is almost irrelevant form the experimental point of view.[1] The critical phenomena were, however, observed in high-T_c superconductors.

Liquid crystals

Liquid crystals consist of long molecules. At the 1st approximation they can be considered as inflexible rods. The simplest phase is *nematic* where the rods are oriented. Otherwise the nematic is a disordered liquid.

There exist more complicated structure like *cholesteric* phase where a director rotates at a constant pitch in a plane perpendicular to the direction in which one moves. The period is about 3000 Å. *Smectic* liquid crystals combine orientation order with a spatial one along *one* direction.

Structural phase transitions

These transitions are changes in the spatial symmetry of the crystalline lattice.

Other phase transitions

One can name many other phenomena which can be regarded as critical ones, e. g. change in the symmetry of adsorbed monolayers, percolation, dynamical phase transitions, etc.

12.1 Landau Theory

Landau theory was designed to describe systems close to a phase transition where the order parameter ψ is small. ψ can be a scalar, vector, etc. Near the transition point we can expand the free energy as

$$\tilde{G}(P,T,\psi) = G_0 + \alpha\psi + A\psi^2 + C\psi^3 + B\psi^4 \ldots .$$

Having in mind the magnetization as an example, we can specify magnetization M as an order parameter, then the free energy depends on T,H,M where H is the external magnetic field. Those

---

[1] The situation is a bit different at low dimensions.

are the *intensive* parameters which characterize the state. Note that G is not a true Gibbs free energy, it is a hypothetical free energy which would be the case if the system with the order parameter ψ could exist as an equilibrium one. The equilibrium is reached at the minimum of G, and in the minimum it coincides with the Gibbs free energy.

The total scheme of the Landau theory is the following.

- The free energy must be properly defined and then expanded in a Taylor series *up to the minimal* relevant order (see below).
- The spontaneous value(s) of the order parameter is found by minimization of $\tilde{G}$ with respect to ψ keeping the thermodynamic parameters constant.

Uniaxial ferromagnets

In zero magnetic field there is up-down symmetry, $\tilde{G}(T,M) = \tilde{G}(T,-M)$. Thus

$$\Delta\tilde{G}(T,M) \equiv \tilde{G}(T,M) - \tilde{G}(T,0) = \frac{r(T)}{2!}M^2 + \frac{u(T)}{4!}M^4 + \cdots .$$

Existence of stable minima requires $u(T)$ to be positive. At $r(T) > 0$ the free energy has one stable minimum, while at $r(T) < 0$ there are 2 equivalent minima $\pm M_O$, see Fig. 12.2. The higher symmetry phase at $r > 0$ is replaced by a lower symmetry phase. This phenomenon is called the *spontaneous symmetry breaking*. At $r(T) = 0$ the high symmetry of the disordered phase in broken and *one* of the low temperature ordered phases is chosen.

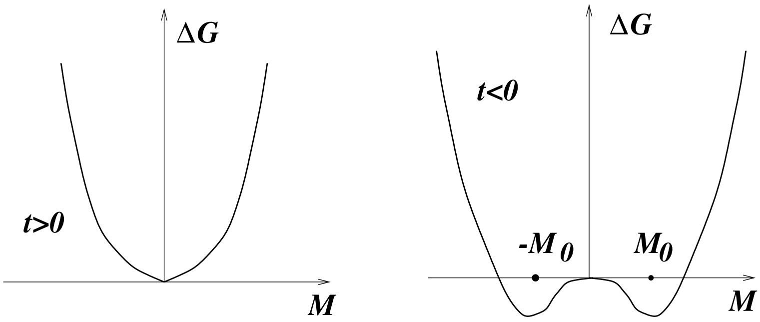

Figure 12.2: The free energy above (left) and below (right) the critical temperature T_c.

Obviously, $r(T_c) = 0$, and near T_c we can assume

$$r(T) = at\,; \quad u(T) = u(T_c) \equiv u\,.$$

Here $\tau = (T - T_c)/T_c$ is the reduced temperature variable. Differentiation of $\tilde{G}$ with respect to M yields

$$0 = \left(\frac{\partial\tilde{G}}{\partial M}\right)_T = a\tau + \frac{u}{3!}M^3\,.$$

This equation has 3 solutions,

$$M = 0\,, \quad M = \pm M_0 = \pm\sqrt{-6a\tau/u}\,.$$

From Fig. 12.2 it is clear that the 1st solution represents a stable minimum above T_c, while the other represent degenerate minima below T_c. If we present the solution as

$$M_0 \propto (-\tau)^\beta \quad \to \psi_0 \propto (-\tau)^\beta$$

we obtain $\beta = 1/2$ for the *critical exponent*. The critical exponents following from the Landau theory are called "classical".

Jump in Specific Heat

Now let us consider the specific heat,

$$C = T\left(\frac{\partial S}{\partial T}\right)_{H=0} = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_{H=0}.$$

Here G is the usual free energy,

$$G(T,H) = \min_M \tilde{G}(T,H,M).$$

At $H = 0$ we get

$$G(T,H) = \begin{cases} G(T,0,0) & ; \quad \tau > 0 \\ G(T,0,M_0(T)) & ; \quad \tau < 0 \end{cases}.$$

As a result, at $\tau < 0$ we obtain

$$G(T,0) = \tilde{G}(T,0,0) + \frac{a\tau}{2}\frac{6a(-\tau)}{u} + \frac{u}{4!}\left(\frac{6a(-\tau)}{u}\right)^2.$$

Now we easily obtain the specific heat by differentiation. Denoting $C_+ = -T\partial^2\tilde{G}(T,0,0)/\partial T^2$ we obtain

$$C = \begin{cases} C_+ & ; \quad \tau > 0 \\ C_+ + (T/T_c)^2 \cdot (3a^2/u) & ; \quad \tau < 0 \end{cases}.$$

Thus we observe a *negative jump* in the specific heat if one passes the critical point from below.

Magnetic Susceptibility

Now let us study a response to a small external magnetic field. Then we have

$$\Delta\tilde{G}(T,M) = -HM + \frac{r(T)}{2!}M^2 + \frac{u(T)}{4!}M^4 + \cdots.$$

The minimization with respect to M leads to the equation for the equilibrium value of the magnetization, M_e,

$$H = rM_e + \frac{1}{3!}M_e^3 + \cdots.$$

The inverse susceptibility is defined as

$$\chi^{-1} = \left(\frac{\partial H}{\partial M_e}\right)_T = r + \frac{u}{2} M_e^2 \,.$$

Substituting the equilibrium magnetization at $H = 0$ we obtain

$$\chi^{-1} = \begin{cases} a\tau & ; \quad \tau > 0 \\ 2a(-\tau) & ; \quad \tau < 0 \end{cases} .$$

If we define the critical exponent of the quantity $A(\tau)$ as

$$s \equiv - \lim_{\tau \to 0_+} \frac{\ln A(\tau)}{\ln \tau}$$

we find the classical critical exponent $\gamma = 1$. We get $\chi(\tau) \propto |\tau|^{-\gamma}$ with the slope ratio $A_+/A_- = 2$.

At the critical isotherm, $\tau = 0$,

$$H = \frac{u}{3!} M_e^3 \quad \to M_e \propto H^{1/3} \,.$$

If we in general assume $M_e \propto H^{1/\delta}$, then the classical value of δ is 3. Later on we shall compare the values of critical exponents with more general theory and with experiment.

Role of External field

Let us now discuss the situation in a finite external field. The plot of the free energy vs. M is shown in Fig. 12.3 for $\tau > 0$ (solid line) and $\tau < 0$ (dashed line). We observe that there is a stable minimum in both cases – magnetic field removed the system from its critical point.

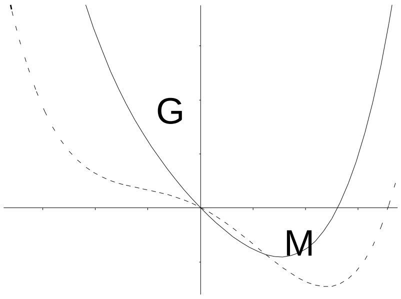

Figure 12.3: The free energy at a finite external field.

Conclusion: If we are interested in continuous phase transition the fields which linearly coupled with the order parameter must be removed.

Role of Cubic Invariants

Now let us consider the case when symmetry of the system allows ψ^3 terms. Then

$$\Delta\tilde{G} = \frac{r(T)}{2!}\psi^2 - \frac{c}{3!}\psi^3 + \frac{u}{4!}\psi^4 + \cdots .$$

Let us put $c > 0$. The analysis of that free energy becomes clear from Fig. 12.4. At sufficiently large positive τ the profile has a shape shown in the left panel. It has one stable minimum at $\psi = 0$. At low temperatures an additional minimum develops, and at some τ_2 this minimum is degenerate with $\psi = 0$ (middle panel). At lower temperatures the second minimum is the only stable one (right panel). Thus there is a (weakly) *first order* transition from $\psi = 0$ to some $\psi = \psi_0$ at $\tau = \tau_2$.

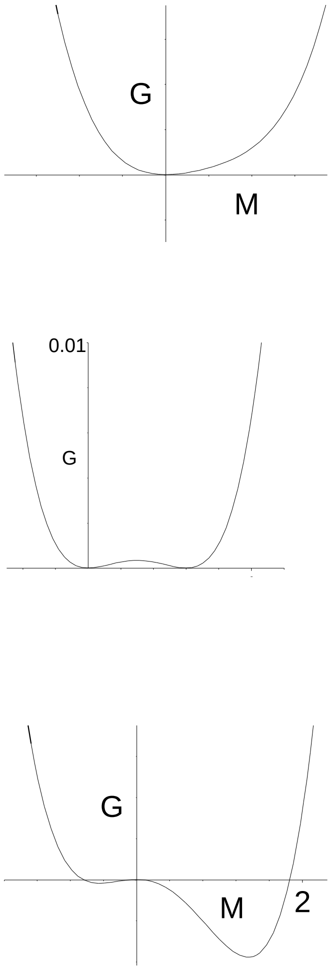

Figure 12.4: The free energy with a cubic invariant included.

Conclusion: Cubic terms results in a first-order transition rather in a continuous one. $\tau = 0$ does not have the significance of the transition temperature in this case.

Tricriticality

Now let us return to the simplest case where the free energy is *even* in the order parameter. Let us assume that $r > 0$, but u is *negative*. To keep the stability one has to add the higher order term to get

$$\Delta\tilde{G} = \frac{r}{2!}\psi^2 + \frac{u}{4!}\psi^4 + \frac{w}{6!}\psi^6 \cdots .$$

At $w > 0$ the system undergoes a 1st order phase transition at some temperature T_0 corresponding to $r = r_0 > 0$. Above T_0 the equilibrium order parameter is zero (see right panel of Fig. 12.5). At $T = T_0$ the transition takes place to one of a symmetric set of states with lower symmetry at which the order parameter is $\pm\psi_0$. The temperature dependence of the order parameter is shown in Fig. 12.6. Suppose now that u is a function of some parameter, p, and it changes sign at $p = p_c$. At this point the character of phase transition will change from 1st order to continuous. The point $r = a\tau = 0,\ u(p_c) = 0$ is called the *tricritical point.* It is easy to find classical critical exponents at the tricritical point. Putting $u = 0$ and minimizing the free energy we obtain the

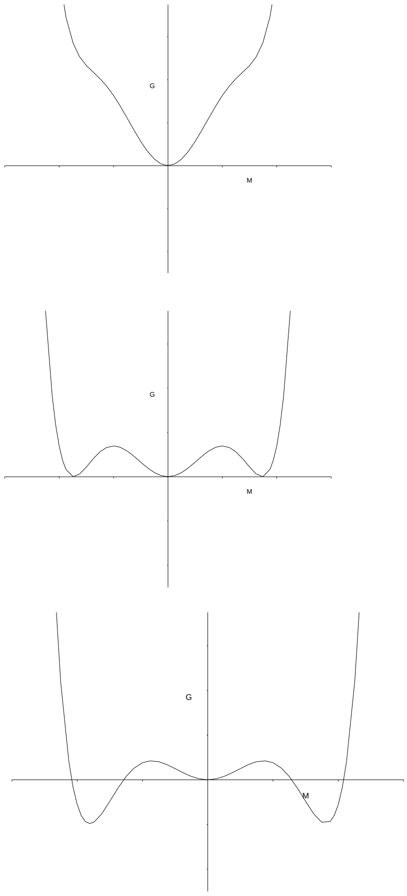

Figure 12.5: The free energy with a cubic invariant included.

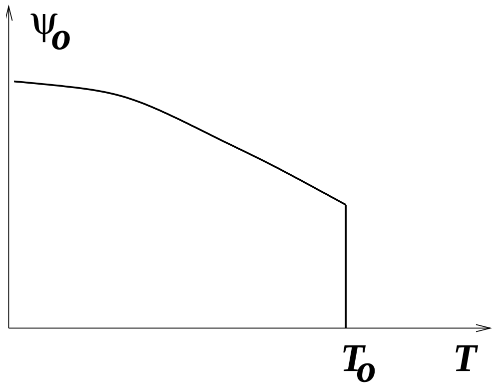

Figure 12.6: The order parameter as a function of the temperature at $u < 0$.

expression for the order parameter,

$$a\tau\psi_0 + \frac{w}{5!}\psi_0^5 = 0 \quad \rightarrow \quad \psi_0 = \left(\frac{5!\,a\,(-\tau)}{w}\right)^{1/4}.$$

Thus $\beta_t = 1/4$.

Summary of the Landau recipe

- *Given* that the system has a continuous or weakly 1st order phase transition form a disordered phase to the phase of higher symmetry.
- Characterize the order of the low symmetry phase and define the order parameter (scalar, vector, tensor, or ...).
- Introduce the proper free energy $\tilde{G}(T,\psi)$ for a given *space-independent* order parameter. The given value of this parameter does not have to be the equilibrium value.
- Write down the Taylor expansion of $\tilde{G}$ in powers of ψ, where all the terms compatible with the symmetries of the system are included. Include only the minimum number of terms.
- Write the coefficients of the ψ-expansion as a Taylor expansion in $\tau = (T - T_c)/T_c$ and keep only the leading ones.[2]
- Determine the equilibrium value(s), ψ_0, by minimizing $\tilde{G}$ with respect to ψ.
- Substitute ψ_0 into the minimum, free energy and calculate equilibrium properties like critical exponents.

---

[2] Sometimes a parameter different from the dimensionless temperature can play a role of τ.

Critique

- The Landau theory *assumes* the existence of the phase transition.
- A general recipe for constructing the order parameter does not exist.
- Spatial dependence of the order parameter is ignored.
- *Fluctuations* of the order parameter are neglected.

12.2 Scaling laws

Consider a system close to continuous phase transition. Let us introduce reduced temperature, τ, and external field h, the transition point corresponding to $(\tau,h)=0$. Since we are interested in *singular* behavior near the critical point let us separate the singular part of the free energy, $G_s(\tau,h)$. The crucial point is to assume (following B. Widom, 1965) that

$$G_s(\lambda^a\tau,\lambda^b h)=\lambda G_s(\tau,h)\,.$$

This is the so-called *homogeneity postulate*. The physical reason for the postulate will be discussed later in connection with renormalization group arguments. The postulate allows one to express all the critical exponents through the constants a and b, and to find relationship between the exponents. These relations are called the *scaling laws*.

To derive scaling laws we choose $\lambda^a=|\tau|^{-1}$. Then,

$$G_s(\tau,h)=|\tau|^{1/a}G_s\left(\frac{\tau}{|\tau|},\frac{h}{|\tau|^{b/a}}\right)\,.$$

Since $\tau/|\tau|=\pm1$ we can express G_s as

$$G_s(\tau,h)=|\tau|^{2-\alpha}\Phi_\pm(h/|\tau|^\Delta)\,,\quad \alpha=2-1/a,\ \Delta=b/a\,.$$

After such an assumption the specific heat is

$$C\sim\frac{\partial^2G_s(\tau,0)}{\partial\tau^2}\sim|\tau|^{-\alpha}$$

while $\Phi_\pm(0)=A_\pm$ are constants. The generalized magnetization is

$$m\sim\frac{\partial G_s}{\partial h}=|\tau|^{2-\alpha-\Delta}\Phi'_\pm(h/|\tau|^\Delta)\,.$$

At $\tau>0$ spontaneous magnetization m_0 vanishes, thus $\Phi'_+(0)=0$. However, when $\tau<0$ $\Phi'_-(0)=B_-\neq0$. As result,

$$\beta=2-\alpha-\Delta\,.$$

In a similar way, the isothermal susceptibility in zero field is given as

$$\chi = \left(\frac{\partial m}{\partial h}\right)_T \sim \frac{\partial^2 G_s}{\partial h^2} = |\tau|^{2-\alpha-2\Delta}\Phi''_\pm(0) .$$

With $\Phi''_\pm(0) = C_\pm \neq 0$ we obtain

$$\gamma = -2+\alpha+2\Delta .$$

Excluding the ”gap exponent” Δ we have

$$\Delta = \beta+\gamma \quad \to \quad \alpha+2\beta+\gamma = 2 .$$

This is the first example of the scaling law. Note that classical exponents are $(\alpha = 0,\ \beta = 1/2,\ \gamma = 1)$ while the exact Ising one for a 2D system are $(\alpha = 0,\ \beta = 1/8,\ \gamma = 7/4)$, as has been derived by L. Onsager. Both sets meet the scaling law.

What is left is to determine the exponent δ for the critical isotherm. The problem is the argument of the function $\Phi'_\pm$ tends to infinity at $|\tau| \to 0$ for finite h. Assuming that $\Phi'_\pm(x) \sim x^\kappa$ for large x we obtain

$$m(|\tau|) \to 0, h) \sim |\tau|^{2-\alpha-\Delta-\kappa\Delta} h^\kappa .$$

For m to have a finite value at $|\tau| \to 0$ we have to put

$$2-\alpha-\Delta-\kappa\Delta \quad \to \quad \delta = \frac{\Delta}{2-\alpha-\Delta} .$$

After a simple algebra we obtain

$$\Delta = \beta\delta = \beta+\gamma \quad \to \quad \alpha+\eta(1+\delta) = 2 ,$$

that is the second scaling law.

Using the scaling laws we can express the behavior of the order parameter as a law of corresponding states as

$$\frac{m}{|\tau|^\beta} = \Phi'_\pm\left(\frac{h}{|\tau|^{\beta\delta}}\right) .$$

12.3 Landau-Ginzburg theory

The Landau-Ginzburg (LG) theory is an attempt to take into account fluctuations of the order parameter in its simplest way. Taking the uniaxial ferromagnet as an example we consider now the *local magnetization* $m(\mathbf{r})$ as the order parameter. Then the free energy becomes a *functional* of $m(\mathbf{r})$ rather the function of a total magnetic moment, M. The recipe is to add only simplest non-trivial terms. Since the free energy must be an even function of magnetization one can imagine the combinations like $m(\mathbf{r})\nabla m(\mathbf{r})$, $m\nabla^2 m$, $(\nabla m(\mathbf{r}))^2$, etc. The 1st combination can be transferred into a surface integral, and the 2nd and the 3d ones can be combined using partial integration. As a result, the LG functional is constructed as

$$G_{LG} = \int d\mathbf{r} \left\{\frac{g}{2}(\nabla m)^2 + \frac{r}{2}m^2(\mathbf{r}) + \frac{u}{4!}m^4(\mathbf{r}) + \cdots\right\} .$$

As before, $r = a\tau,\ u > 0,\ g > 0$. The latter inequality implies that the homogeneous state is the equilibrium one.

The LG theory allows for *long-range* fluctuations because it is an expansion in powers of the magnetization gradient.

Classical fluctuations

In the classical liming discussed we keep only the terms quadratic in fluctuations. Then the fluctuations are *Gaussian* and various fluctuation modes are statistically independent. At $T > T_c$, and with $h = 0$, the spontaneous magnetization $\langle m(\mathbf{r})\rangle = 0$. Consequently, $m(\mathbf{r})$ is itself a fluctuation. To order of m^2 we have

$$G_{LG} = \int d\mathbf{r} \left\{ \frac{g}{2}(\nabla m)^2 + \frac{r}{2}m^2(\mathbf{r}) + \cdots \right\}. \tag{12.1}$$

Below T_c the fluctuation is $\delta m(\mathbf{r}) = m(\mathbf{r}) - m_0(T)$ where $m_0(T) = \langle m(\mathbf{r})\rangle$ is the spontaneous magnetization. The expansion up to $(\delta m)^4$ gives

$$\begin{aligned} G_{LG} \;=\; & \int d\mathbf{r} \left\{ \frac{g}{2}(\nabla\,\delta m)^2 + \frac{r}{2}m_0^2 + \frac{u}{4!}m_0^4 \right. \\ & + \left[r + \frac{u}{3!}m_0^3\right](\delta m) + \left[\frac{r}{2} + \frac{u}{4}m_0^2\right](\delta m)^2 \\ & \left. + \frac{1}{3!}um_0\,(\delta m)^3 + \frac{1}{4!}(\delta m)^4 + \cdots \right\}. \end{aligned} \tag{12.2}$$

According to the Gaussian approximation, the items proportional to $(\delta m)^3$ and $(\delta m)^4$ are neglected. Furthermore, $m_0 = -6r/u$, and the item proportional to δm vanishes. As a result,

$$G_{LG} = \int d\mathbf{r} \left\{ \frac{g}{2}(\nabla\,\delta m)^2 + \frac{(-2r)}{2}m_0^2 + \cdots \right\}. \tag{12.3}$$

Since

$$\int d\mathbf{r}\, A(\mathbf{r})B(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{k}} A_{-\mathbf{k}}B_{\mathbf{k}} \to \int \frac{d\mathbf{k}}{(2\pi)^d} A_{-\mathbf{k}}B_{\mathbf{k}}$$

we obtain above T_c

$$G_{LG} = \frac{1}{2V}\sum_{\mathbf{k}}{}' (gk^2 + r)m_{-\mathbf{k}}m_k. \tag{12.4}$$

$\sum'$ means the we restrict summation up to some k_0 we have to do that because we neglect higher gradients.

Now we are able to calculate correlation functions like

$$\Gamma^m_{\mathbf{q}\mathbf{q}'} \equiv \langle m_{\mathbf{q}} m_{\mathbf{q}'}\rangle = \frac{\int \prod'_{\mathbf{k}} dm_{\mathbf{k}}\, m_{\mathbf{q}} m_{\mathbf{q}'} \exp(-G_{LG})}{\int \prod'_{\mathbf{k}} dm_{\mathbf{k}} \exp(-G_{LG})}.$$

Here $\prod'_{\mathbf{k}}$ means that one has to take into account only $|\mathbf{k}| \le k_0$. Since $m(\mathbf{r})$ is real, $m_{\mathbf{k}} = m^*_{-\mathbf{k}}$, it is clear from(12.4) that

$$\Gamma^m_{\mathbf{q}\mathbf{q}'} = V\,\delta_{\mathbf{q}',-\mathbf{q}}\,\Gamma^m_{\mathbf{q}}$$

where

$$\Gamma^m_{\mathbf{q}} \equiv \langle m_{-\mathbf{q}} m_{\mathbf{q}}\rangle = \frac{\int \prod'_{\mathbf{k}} dm_{\mathbf{k}}\, m_{-\mathbf{q}} m_{\mathbf{q}} \exp(-G_{LG})}{\int \prod'_{\mathbf{k}} dm_{\mathbf{k}} \exp(-G_{LG})}\,.$$

Let us count only those $\mathbf{k}$ for which $k_x > 0$. Then

$$\begin{aligned}
\prod_{\mathbf{k}}{}' dm_{\mathbf{k}} &= \prod_{\mathbf{k}}{}'' dm_{-\mathbf{k}} dm_{\mathbf{k}} \\
\sum_{\mathbf{k}}{}' (gk^2 + r)\, m_{-\mathbf{k}} m_{\mathbf{k}} &= 2\sum_{\mathbf{k}}{}'' (gk^2 + r)\, m_{-\mathbf{k}} m_{\mathbf{k}}\,.
\end{aligned}$$

As a result, we are left with the integral

$$\Gamma^m_{\mathbf{q}} = \frac{1}{V}\frac{\int dm_{-\mathbf{q}} dm_{\mathbf{q}}\, m_{-\mathbf{q}} m_{\mathbf{q}} \exp[-(gq^2 + r) m_{-\mathbf{q}} m_{\mathbf{q}}/V]}{\int dm_{-\mathbf{q}} dm_{\mathbf{q}} \exp[-(gq^2 + r) m_{-\mathbf{q}} m_{\mathbf{q}}/V]}\,.$$

Since $m_{-\mathbf{q}}$ is a complex variable we can parameterize it as $m_{\mathbf{q}} = me^{i\phi}$ and calculate the Jacobian

$$\frac{\partial(m_{-\mathbf{q}}, m_{\mathbf{q}})}{\partial(m,\phi)} = \begin{vmatrix} e^{-i\phi} & e^{i\phi} \\ -ime^{-i\phi} & ime^{i\phi} \end{vmatrix} = 2im\,.$$

Then we can introduce a new variable $x = m^2$, $dx = 2m\,dm$ and use the result for Gauss integral. As a result,

$$\Gamma^m_{\mathbf{q}} = \frac{1}{gq^2 + a\tau}\,.$$

This formula has been obtained previously by Ornstein and Zernike in 1914 using other considerations. After Fourier transform to real coordinates we obtain for $d > 1$ and $r \gg \xi$

$$\Gamma^m(r) \sim \begin{cases} r^{(1-d)/2} e^{-r/\xi} & \text{for} \quad r \gg \xi,\ d > 1 \\ r^{2-d} & \text{for} \quad r \ll \xi,\ d > 2 \end{cases} \quad ; \quad \xi = \sqrt{g/a\tau}$$

In this way we predict critical exponents for the correlation length, $\nu = 1/2$.

The Ginzburg criterion

Let us estimate the validity of the approximation made. The measure of the total strength of the long-wave fluctuations are given by the integral

$$\int d\mathbf{r}\, \Gamma^m(\mathbf{r}) \equiv \Gamma^m_0\,, \quad \Gamma^m_0 = \Gamma^m_{\mathbf{q}=0} = (a\tau)^{-1}\,.$$

Let us denote a typical value of the fluctuation squared by $\overline{(\delta m)^2}$. Since the typical diameter of correlated region is ξ we have

$$\Gamma_0^m \approx \overline{(\delta m)^2}\,\xi^d\,.$$

Consequently,

$$\overline{(\delta m)^2} \sim \xi^{-d}\,(a\tau)^{-1} \sim \frac{1}{a}\left(\frac{a}{g}\right)^{d/2}\tau^{(d/2)-1}\,.$$

Now we can estimate when we can neglect 4th order terms in the expansion in powers of fluctuations. A typical value of m when the 4th and 2nd order terms are comparable is

$$\frac{r}{2}\bar{m}^2 = \frac{u}{4!}\bar{m}^4 \quad \to \quad \bar{m}^2 = \frac{12a\tau}{u}\,.$$

The LG theory is valid at $\overline{(\delta m)^2} \ll \bar{m}^2$, or at

$$\tau^{(d-4)/2} \ll \frac{a^2}{u}\left(\frac{g}{a}\right)^{d/2}\,.$$

This is the famous Ginzburg criterion. We see that at $d > 4$ this criterion is always satisfied because the r.h.s. is a constant depending on the concrete problem under consideration. Thus $d_c = 4$ is called the *upper critical dimensionality*. Accordingly, *all the critical exponents* take their classical values.

At $d < d_c = 4$ the Ginzburg criterion is violated close to the critical point. For $d = 3$ the criterion has the form

$$\tau \gg u^2/sg^3\,,$$

so the LG theory is reasonable outside that region. The upper critical dimensionality $d_c = 4$ is valid for all the universality classes (d,n). We postpone a discussion what is going on in a close vicinity of the critical point.

12.4 Ginzburg-Landau theory of superconductivity (SC)

Let us formulate the theory of SC transition as the one for a general continuous transition. First, one should define the order parameter. As the order parameter it is natural to chose the *wave function of the Bose condensate*, Ψ. According to the principles of SC, it is natural to assume that the ground state corresponds to the total momentum equal to 0, the wave function being the same for all the Bose particles.

Near the critical temperature, the modulus of the order parameter is small, and one can expand the thermodynamic potential as

$$\mathcal{G}_s = \mathcal{G}_n + a|\Psi|^2 + \frac{b}{2}|\Psi|^4 + \cdots\,.$$

Because we are interested in the vicinity of T_c we can expand the coefficients in power of

$$\tau = \frac{T - T_c}{T_c}.$$

According the definition of T_c, above T_c the stable state corresponds to $\Psi = 0$, while below T_c $\Psi \neq 0$. Thus the coefficient a should change the sign at the transition point:

$$a = \alpha\tau, \quad \alpha > 0.$$

As a result, we describe the equilibrium values of the order parameter as

$$\begin{array}{ll} \Psi = 0 & \text{at} \quad T > T_c, \\ |\Psi|^2 = -(\alpha/b)\tau = |\Psi_0|^2 & \text{at} \quad T < T_c \end{array}.$$

Inserting the equilibrium value if $|\Psi_0|^2$ into the expansion and comparing the result with the known relation for the critical field, we get

$$\mathcal{G}_s - \mathcal{G}_n = \frac{(\alpha\tau)^2}{2b} = \frac{H_c^2}{8\pi}.$$

The last equality is a consequence of the definition of the critical field. Indeed $H_c^2/8\pi$ is just the energy of magnetic field repelled from the sample in the Meißner state. The ratio $\alpha^2/2b$ can be related to the critical temperature T_c from the BCS theory of superconductivity. [3] The most interesting case is the case of external magnetic field. In this case Ψ is *coordinate dependent*. Thus we should add both the energy of magnetic field $H^2/8\pi$ and the energy connected with the inhomogeneity. Near the critical point it is enough to add $|\nabla\Psi|^2$. Here we come to the most important point: Cooper pairs are *charged particles*. Consequently, because of gauge invariance, only combination

$$-i\hbar\nabla + \frac{2e}{c}\mathbf{A}$$

is possible. To make a proper dimension we write the corresponding term as

$$\frac{1}{4m}\left|\left(-i\hbar\nabla + \frac{2e}{c}\mathbf{A}\right)\Psi\right|^2.$$

Finally, we get for $\delta\mathcal{G} = \mathcal{G}_s - \mathcal{G}_n^0$

$$\delta\mathcal{G} = \int dV \left\{\alpha\tau|\Psi|^2 + \frac{b}{2}|\Psi|^4 + \frac{1}{4m}\left|\left(-i\hbar\nabla + \frac{2e}{c}\mathbf{A}\right)\Psi\right|^2 + \frac{H^2}{8\pi}\right\}.$$

Here $\mathcal{G}_n^0$ is the free energy of the normal state *without magnetic field.* To get the minimum we calculate variation with respect to Ψ^*:

$$\int d\mathcal{V}\left\{\alpha\tau\Psi\delta\Psi^* + b|\Psi|^2\Psi\delta\Psi^* + \frac{1}{4m}\left(-i\hbar\nabla + \frac{2e}{c}\mathbf{A}\right)\Psi\left(i\hbar\nabla + \frac{2e}{c}\mathbf{A}\right)\delta\Psi^*\right\}$$

---

[3] $\alpha^2/2b = c1g(\varepsilon_F)(k_BT_c)^2$. where $g(\varepsilon_F)$ is the density of states at the Fermi level while c_1 is a numerical constant.

The item with the $\nabla\delta\Psi^*$ can be integrated by parts. Using the Gauss theorem we get

$$\frac{i\hbar}{4m}\oint_S \delta\Psi^*\left(-i\hbar\nabla+\frac{2e}{c}\mathbf{A}\right)\Psi\,d\mathbf{S}+\frac{1}{4m}\int d\mathcal{V}\,\delta\Psi^*\left(-i\hbar\nabla+\frac{2e}{c}\mathbf{A}\right)^2\Psi.$$

Then we put $\delta\Psi^*=0$ at the surface and arbitrary inside. As a result, we get the following equation for the order parameter

$$\frac{1}{4m}\left(-i\hbar\nabla+\frac{2e}{c}\mathbf{A}\right)^2\Psi+\alpha\tau\Psi+b|\Psi|^2\Psi=0.$$

If now we put $\delta\Psi^*$ at the surface to be arbitrary, we obtain the boundary condition

$$\mathbf{n}\left(-i\hbar\nabla+\frac{2e}{c}\mathbf{A}\right)\Psi|_S=0.$$

The variation with respect to $\delta\Psi$ leads to the complex conjugate expressions.

Now it is important to get the equation for electric current. To obtain it one should calculate the variation with respect to the vector potential $\mathbf{A}$. The variation of H^2 leads to

$$\delta(\text{curl }\mathbf{A})^2=2\,\text{curl }\mathbf{A}\,\text{curl }\delta\mathbf{A}.$$

Then we can use the relation

$$\text{div }[\mathbf{a}\times\mathbf{b}]=\mathbf{b}\,\text{curl }\mathbf{a}-\mathbf{a}\,\text{curl }\mathbf{b}$$

to get

$$\delta(\text{curl }\mathbf{A})^2=2\delta\mathbf{A}\,\text{curl curl }\mathbf{A}+2\text{ div }[\delta\mathbf{A}\times\text{curl }\mathbf{A}].$$

The second integral can be transformed into the surface one, thus it is not important. On the other hand, from the Maxwell equations

$$\text{curl curl }\mathbf{A}=\text{curl }\mathbf{H}=\frac{4\pi}{c}\mathbf{j}\quad\rightarrow\quad\delta(\text{curl }\mathbf{A})^2=2\mathbf{j}\delta\mathbf{A}.$$

The rest of the free energy can be calculated in a straightforward way, and equating the total variation to zero we obtain

$$\mathbf{j}=\frac{ie\hbar}{2m}(\Psi^*\nabla\Psi-\Psi\nabla\Psi^*)-\frac{2e^2}{mc}|\Psi|^2\mathbf{A}$$

This is just the quantum mechanical expression for the current for a particle with the charge $(-2e)$ and the mass $2m$. The set

$$\begin{gathered}(1/4m)\left[-i\hbar\nabla+(2e/c)\mathbf{A}\right]^2\Psi+\alpha\tau\Psi+b|\Psi|^2\Psi=0,\\ \mathbf{j}=(ie\hbar/2m)(\Psi^*\nabla\Psi-\Psi\nabla\Psi^*)-(2e^2/mc)|\Psi|^2\mathbf{A}\end{gathered}$$

forms the *Ginzburg-Landau equations.*

Dimensionless parameters and their meaning

To make the equations simples new dimensionless parameters are usually introduced:

$$\Psi' = \frac{\Psi}{\Psi_0}, \qquad \mathbf{H}' = \frac{\mathbf{H}}{H_c\sqrt{2}}, \qquad \mathbf{r}' = \frac{\mathbf{r}}{\delta}, \qquad \mathbf{A}' = \frac{\mathbf{A}}{H_c\delta\sqrt{2}}$$

$$\text{with} \quad \delta = \sqrt{\frac{2mc^2}{4\pi(2e)^2\Psi_0^2}}, \qquad H_c = 2\sqrt{\pi}\frac{\alpha\tau}{\sqrt{b}}, \qquad \Psi_0^2 = \frac{\alpha|\tau|}{b}.$$

The quantity Ψ_0 is a spatially-homogeneous solution of the GL equations in the absence of external magnetic field. In the following we will omit primes and use these variables. As a result, the set of equations can be expressed as

$$\begin{aligned} \left(\frac{i}{\kappa}\nabla - \mathbf{A}\right)^2 \Psi - \Psi + |\Psi|^2\Psi &= 0, \\ \text{curl curl}\mathbf{A} - \frac{i}{2\kappa}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) - |\Psi|^2\mathbf{A} &= 0 \end{aligned}$$

with the boundary condition

$$\mathbf{n}\left(\frac{i}{\kappa}\nabla - \mathbf{A}\right)\Psi_{\text{surface}} = 0.$$

This set contains only one dimensionless parameter

$$\kappa = 2\sqrt{2}\frac{eH_c\delta^2}{\hbar c}$$

which is called the *Ginzburg-Landau*(GL) *parameter.*

To understand the meaning of the parameters let us consider the simplest problem, namely penetration of a weak field into the superconductor $x > 0$. Let $\mathbf{H} \parallel \mathbf{z}$, $\mathbf{A} \parallel \mathbf{y}$, (see Fig. 12.7). We have

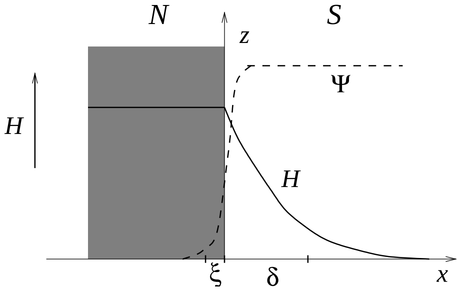

Figure 12.7: On the magnetic field penetration into a superconductor. Solid line is the field profile while dashed one shows spatial dependence of the order parameter.

$$H(x) = \frac{dA_y}{dx}.$$

It is natural to consider Ψ as a function only of x. We get

$$-\frac{1}{\kappa^2}\frac{d^2\Psi}{dx^2} - \Psi + A^2|\Psi| = 0, \qquad \left.\frac{d\Psi}{dx}\right|_{\text{surface}} = 0.$$

At $\kappa \ll 1$ one can assume Ψ =const, in the bulk of material it is natural to assume $|\Psi|^2 = 1$, we chose Ψ real. Immediately we get from the second GL equation

$$\frac{d^2A}{dx^2} - A = 0 \quad \to \quad H = H_0\, e^{-x}, \quad \text{in usual units} \quad H = H_0\, e^{-x/\delta}\,.$$

Thus δ is just the London penetration depth near T_c.

Application of the GL theory

Surface energy of N-S interface

First, let us derive an auxiliary expression for the free energy which will be used below. Using dimensionless notations we get

$$\int \left(\Omega_s - \Omega_n^{(0)}\right) d\mathcal{V} = \frac{H_c^2}{4\pi} \int d\mathcal{V} \left\{ -|\Psi|^2 + \frac{|\Psi|^4}{2} + \left| \left(\frac{i\nabla}{\kappa} - \mathbf{A} \right) \Psi \right|^2 + H^2 \right\}.$$

Then we integrate the item with $\nabla\Psi^*$ by parts, the surface integral being zero. In the rest integral we take into account that Ψ obeys the GL equation. As a result,

$$\int \left(\Omega_s - \Omega_n^{(0)}\right) d\mathcal{V} = \frac{H_c^2}{4\pi} \int d\mathcal{V} \left[H^2 - \frac{|\Psi|^4}{2} \right].$$

If one keeps the external field $\mathbf{H}_0$ fixed the corresponding thermodynamic potential can be obtained by subtracting $\mathbf{H}_0\mathbf{B}/4\pi$. In a normal phase, correspondingly $\Omega_{nH} = \Omega_n^{(0)} - H_0^2/8\pi$. As a result,

$$\int \left(\Omega_{sH} - \Omega_{nH}\right) d\mathcal{V} = \frac{H_c^2}{4\pi} \int d\mathcal{V} \left[(H - H_0)^2 - \frac{|\Psi|^4}{2} \right]. \tag{12.5}$$

Consider the problem where all the quantities depend only upon one co-ordinate, say x, while $\mathbf{A} \perp \mathbf{x}$. The set of equations has the form (we'll see that Ψ is *real*)

$$\begin{aligned} \frac{1}{\kappa^2}\frac{d^2\Psi}{dx^2} + \Psi(1 - A^2) - \Psi^3 &= 0, \quad \frac{d\Psi}{dx}\Big|_{\text{surface}} = 0\,, \\ \frac{d^2A}{dx^2} - \Psi^2 A &= 0. \end{aligned}$$

It is easy to find the first integral of this set. Let us multiply the first equation by $d\Psi/dx$, the last one – by dA/dx and them add the equation and integrate over x:

$$\int_0^x dx \left[\frac{1}{\kappa^2}\frac{d^2\Psi}{dx^2}\frac{d\Psi}{dx} + \Psi\frac{d\Psi}{dx}(1 - A^2) - \frac{d\Psi}{dx}\Psi^3 + \frac{d^2A}{dx^2}\frac{dA}{dx} - \Psi^2 A\frac{dA}{dx} \right] = 0.$$

We obtain

$$\frac{1}{\kappa^2}\left(\frac{d\Psi}{dx}\right)^2 + \Psi^2(1 - A^2) - \frac{\Psi^4}{2} + \left(\frac{dA}{dx}\right)^2 = \text{const} = \frac{1}{2}$$

(the constant being determined from the boundary conditions).

To find the surface energy we formulate the boundary conditions

$$\begin{array}{lllll} x\to\infty & \text{(S):} & \Psi=1, & H=A=0, & d\Psi/dx=0, \\ x\to-\infty & \text{(N):} & \Psi=0, & H=H_0=1/\sqrt{2}, & d\Psi/dx=0. \end{array}$$

At $\kappa \ll 1$ the most important region where A and H are small. Thus, we get

$$\frac{1}{\kappa^2}\left(\frac{d\Psi}{dx}\right)^2+\Psi^2-\frac{\Psi^4}{2}=\frac{1}{2} \quad \to \quad \frac{d\Psi}{dx}=\frac{\kappa}{\sqrt{2}}(1-\Psi^2).$$

The solution is

$$\Psi=\tanh\frac{\kappa x}{\sqrt{2}},$$

it is wrong in the region where the field penetrates, but this region is small. We observe that the typical length at which the order parameter is changed is (in dimensional units) $\xi=\kappa\delta$. ξ is called *the coherence length*.

Now we can employ Eq. (12.5) and put $H_0=1/\sqrt{2}$, $H=0$. We obtain

$$\sigma_{ns}=\frac{H_c^2}{8\pi}\int\left[1-\tanh^2\frac{\kappa x}{\sqrt{2}}\right]dx=\frac{H_c^2}{8\pi}\frac{4\sqrt{2}}{3\kappa}.$$

In dimensional units that means

$$\sigma_{ns}=\frac{4\sqrt{2}}{3}\frac{H_c^2}{8\pi}\frac{\delta}{\kappa}.$$

The opposite limiting case can be treated only numerically. The result is that $\sigma_{ns}=0$ at $\kappa=1/\sqrt{2}$ and at $\kappa>1/\sqrt{2}$ it is negative. This property has very important implication on the magnetic field structure in superconductors.

Quantization of magnetic flux

Consider a hollow cylinder placed into a longitudinal magnetic field. Assume that the cylinder is in the Meißner state. If the order parameter has the form

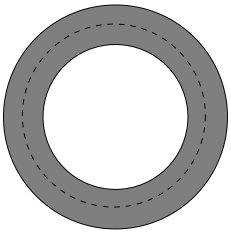

Figure 12.8: On the flux quantization in a hollow cylinder

$$\Psi=|\Psi|e^{i\chi}$$

the current is

$$\mathbf{j} = -\frac{e\hbar}{m}|\Psi|^2 \left[\nabla\chi + \frac{2e}{\hbar c}\mathbf{A}\right].$$

We can divide the current density by $|\Psi|$ and integrate over a closed circuit in the bulk of superconductor:

$$\underbrace{\oint \frac{\mathbf{j}\cdot d\mathbf{l}}{|\Psi|^2}}_{0} = -\frac{e\hbar}{m}\left[\underbrace{\oint d\mathbf{l}\cdot\nabla\chi}_{-2\pi k} + \frac{2e}{\hbar c}\underbrace{\oint \mathbf{A}\, d\mathbf{l}}_{\text{flux}}\right].$$

Thus

$$\Phi = k\Phi_0\,, \quad \Phi_0 = \pi\hbar c/e = 2.07\cdot 10^{-7}\ \mathrm{G}\cdot\mathrm{cm}^2.$$

It is interesting that the effect has been predicted by F. London (1950) before the concept of Cooper pairs. He assumed the quantum to be $\pi\hbar c/e = 2\Phi_0$.

In any case the quantized quantity is in fact *fluxoid*

$$\oint \nabla\chi\, d\mathbf{l}$$

which is equal to the total flux through the hole and the surface layer. As a result, the flux quantization through the hole is not exact.

Quantization of magnetic flux through a hollow cylinder has been experimentally observed and has many implications.

12.5 Mean field theory

The mean field approximation

This is a basic approximation which allows one to find main results regarding phase transitions leaving alone some subtle problems about correct critical exponents. For simplicity, let us restrict ourselves with the generic Ising model with the partition function

$$Z_N = \sum_s \exp\left(\beta\sum_{i<j} J_{ij}s_is_j + \beta\sum_i h_is_i\right).$$

The main idea is to look at the world from a point of view of a given spin. A given spin i "feels" the external field + the fields created by the neighbors,

$$h_i(\{s\}) = \sum_{ij}{}' J_{ij}\,s_j + h_i\,,$$

where prime means that $i \neq j$. The first step is to replace the actual field by the average one,

$$h_i(\{s\}) \to h_i^e = \sum_{ij}{}' J_{ij}\,\langle s_j\rangle + h_i = \sum_{ij}{}' J_{ij}\,m_i + h_i\,.$$

Then

$$Z_N \to Z^e = \sum_s e^{\beta \sum_i h_i^e s_i} = \prod_{i=1}^N 2\cosh(\beta h_i^e)\,.$$

Then we have to close the scheme, i. e. to calculate $m_i = \langle s_i \rangle$. Since

$$m_i = \frac{1}{\beta}\frac{\partial \ln Z_N^e}{\partial h_i^e} = \tanh(\beta h_i^e)$$

we have the equation

$$m_i = \tanh\left[\beta\left(\sum_j{}' J_{ij} m_j + h_i\right)\right]$$

which should be solved for m_i.

In principle, it is a set of many equations, and further simplifications are required. Before doing that we shall list some comments.

- The mean field approximation neglects the fluctuations of the order parameter. So the critical exponents are classical.
- Contrary to the Landau theory, the mean field theory can be used even when the order parameter is not small. This is why it can be used to discuss *global properties* of the phase diagram.
- The mean field theory works better if the interaction range is large.

12.6 Spatially homogeneous ferromagnets

Assume the spatial homogeneity, $h_i = h$, and that surface effects are not important. Then $J_{ij} = J(\mathbf{r}_i, \mathbf{r}_j) = J(|\mathbf{r}_i - \mathbf{r}_j|)$. We obtain

$$m = \tanh\left[\beta\left(\bar{J}m + h\right)\right]\,, \quad \bar{J} = \sum_j{}' J_{ij}\,.$$

Let us discuss the case $h = 0$. The situation is illustrated in Fig. 12.9 where the l.h.s. and r.h.s. are plotted At small β (high temperatures) there is a unique solution $m = 0$, while at low T (large β) 3 solutions exist. Analysis of the free energy shows that the solution $m = 0$ is unstable, the other ones are degenerate at $h = 0$. At the critical point all 3 solutions merge. That takes place at

$$\left(\frac{\partial \tanh(\beta \bar{J} m)}{\partial m}\right)_{m=0} = \beta \bar{J} = 1\,, \quad \to \quad T_c = \bar{J} = \sum_j{}' J_{ij}\,.$$

With nearest neighbor interaction and for a (hyper)cubic lattice $\bar{J} = 2dJ$ where J is the nearest neighbor interaction. [4]

---

[4] At $d = 1$ the predictions of the mean field theory are wrong even *qualitatively*.

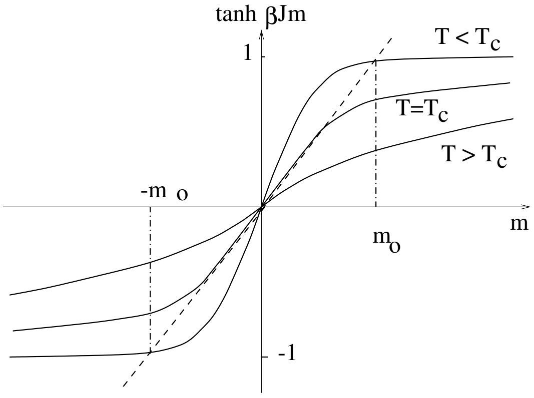

Figure 12.9: Graphical solution of the mean field equation.

Let us re-derive the same expression form another point of view which seems useful for following discussions. From the mean field equation we we can calculate the susceptibility

$$\chi = \left(\frac{\partial m}{\partial h}\right)_T = \cosh^{-1}\left[\beta\,(\bar{J}m + h)\right]\left[\beta\bar{J}\left(\frac{\partial m}{\partial h}\right)_T + \beta\right].$$

Putting $h = 0$ and having in mind that at $T > T_c$ $m = 0$ we obtain

$$\chi = \frac{\beta}{1-\beta J} = \frac{T}{T-T_c}.$$

In this way we obtain that susceptibility id divergent at the critical point, and the critical exponent $\gamma = 1$.

12.7 Renormalization framework

The main idea (L. Kadanoff, 1966) is that since $\xi \to \infty$ at the critical point there is no intrinsic scale for the long range fluctuations. Thus the fluctuations must have the similar appearance at all the scales. Geometrically speaking, the critical state should have properties of a fractal. Thus, *critical Hamiltonian* should exist for the variables on any scale. The Kadanoff's idea was a starting point for K. Wilson who has developed a quantitative approach (Nobel Prize, 1983). The method is a *mapping*, or the *renormalization group transformation* which transforms the Hamiltonian from one scale into one describing the system on a larger length scale.

The proper way to perform such a scaling is to keep long-range (critical) fluctuations unchanged. We shall briefly discuss only one way to perform such a scaling – the *real space renormalization* constructed by Th. Niemeijer and J.M.K. van Leeuven.

The mapping

Let us consider a general class of Ising models. The N spins take the values $s_i = \pm 1$, the Hamiltonian can be written as

$$\mathcal{H}(s) = \sum_i K_i s_i + \sum_{ik} K_{ik}\, s_i s_k + \sum_{ikl} K_{ikl}\, s_i s_k s_l + \cdots \equiv \sum_\alpha K_\alpha \Pi_\alpha(s). \qquad (12.6)$$

Let us fix the origin on the energy scale by the stipulation

$$\sum_s \mathcal{H}(s) = 0\,.$$

The equation (12.6) can be inverted to express the coupling constants K_β as[5]

$$K_\beta = \frac{1}{2^N}\sum_s \Pi_\beta(s)\mathcal{H}(s) \tag{12.7}$$

The thermodynamics of the model is contained in the partition function,

$$Z_N = \sum_s e^{-\beta\mathcal{H}(s)}\,, \quad \mathcal{G} = -T\lim_{N\to\infty}\frac{1}{N}\ln Z_N\,.$$

Here $\mathcal{G}$ is the free energy per spin.

Our goal is to map the effective Hamiltonian (12.6) to another one relevant to a larger scale. Consider spins on a 2D triangular lattice, as shown in Fig. 12.10. Let us consider a triangle as a new spin. The "coarse-grained" spins are defined according to the algorithm that it is the same as the majority of spins in the covered triangle. This mapping can be analytically expressed as a projection operator,

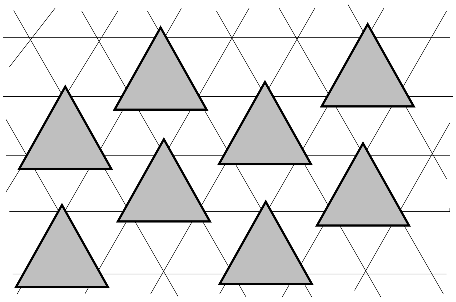

Figure 12.10: Cells in a triangular lattice.

$$P(s',s) = \prod_i P_i\,, \quad P_i = \left\{ \begin{array}{ll} 1 & \text{if } s'_i = \text{sgn}\,(s^i_1 + s^i_2 + s^i_3) \\ 0 & \text{otherwise} \end{array} \right.\,.$$

Here i labels the cell while s^i_k are original spins in the cell i. Then we can write,

$$Z_N = \sum_{s'}\sum_s P(s',s)e^{-\beta\mathcal{H}(s)}\,, \quad \sum_{s'} P(s',s) = 1\,.$$

Now we can determine a new Hamiltonian which is valid on the new scale as

$$\sum_s P(s',s)e^{-\beta\mathcal{H}(s)} \equiv e^{\beta\Omega_1 - \beta\mathcal{H}'(s')}\,.$$

---

[5] To prove this equation one can substitute (12.6) into (12.7). For every pair of sets $\alpha \neq \beta$ the spin s_k summed over all the states gives zero, $\sum_k s_k = 0$. When $\beta = \alpha$ each term gives unity, and there are 2^N such terms.

We have to introduce a spin-independent exponent $\beta\Omega_1$ to keep the property

$$\sum_{s'} \mathcal{H}'(s') = 0\,.$$

This is a unique definition, and using the inversion scheme (12.7) we can determine a new set k'_α of the coupling constants. So we know, at least in principle, the mapping $\mathbf{K} \to \mathbf{K}'(\mathbf{K})$.

The mapping $\mathcal{H}(s) \to \mathcal{H}(s')$ is called the *renormalization group transformation* (RG). In this way we map a triangular lattice on another triangular lattice increasing the scale by $\sqrt{3}$. So we can expect that singularities are the same for both Hamiltonians.

Global properties of the mapping

As a result of RG transform we get

$$Z_N = \sum_{s',s} P(s',s)\, e^{-\beta\mathcal{H}(s)} = e^{\beta\Omega_1} Z'_{N'}\,.$$

Here $Z'_{N'}$ is the partition function for the transformed Hamiltonian with N' cells. Since $N'/N = l^{-d}$ where l is the scaling factor,

$$\mathcal{G}(\mathbf{K}) = \mathbf{\Omega}_1 + \frac{1}{l^d}\mathcal{G}(\mathbf{K}')\,, \quad \mathbf{\Omega}_1 = \Omega_1/N\,.$$

The spin-independent item is calculated from the initial configuration when the fluctuations are frozen. Thus it is not responsible for the singular behavior. Keeping only singular items we require

$$\mathcal{G}(\mathbf{K}) = \frac{1}{l^d}\mathcal{G}(\mathbf{K}')\,.$$

Before using the RG scheme let us recall that the physical correlation length remains the same. However while rescaling we use different yard sticks, so $\xi(\mathbf{K}') = \frac{1}{l}\xi(\mathbf{K})$. Since at the critical point $\xi(\mathbf{K}_c) = \infty$ it follows that the critical Hamiltonian is mapped also on the another critical Hamiltonian.

To demonstrate the procedure, let us choose the 2D set of the coupling constants, $\mathbf{K} = (K_1, K_2)$. An example is a magnetic system in a zero magnetic field having nearest and next-nearest neighbor interaction. To discuss dynamics let us recall that the quantities $\mathbf{K}^e = -\beta\mathbf{K}$ rather than $\mathbf{K}$ enter the the theory. Thus changing the temperature we can move radially in the plane of effective coupling constants $\mathbf{K}^e$. The origin of the (K_1^e, K_2^e)-plane corresponds to infinite temperature, while $T = 0$ corresponds to the infinity in the direction given by the ratio K_2/K_1 (dashed line in Fig. 12.11). Since for every such direction there is a critical temperature, the *subspace of the critical Hamiltonians* is represented by a *line of criticality* in the (K_1^e, K_2^e)-plane.

Suppose that we start from a *supercritical* Hamiltonian represented by point 1. The repeated RG transforms generate Hamiltonians 2, 3, 4, $\cdots$. The correlation length of the initial Hamiltonian was finite, so in shrinks by l after each iteration. Thus the Hamiltonian moves more and more away from criticality. Eventually the point $(K_1^e, K_2^e) = (0,0)$ is reached. In other words, the

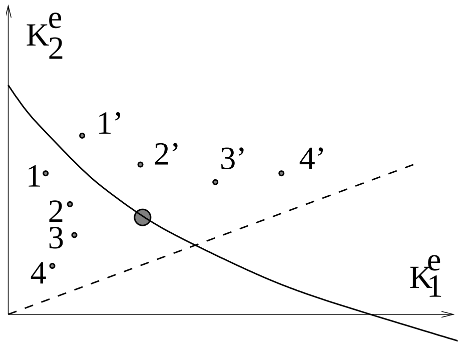

Figure 12.11: Results of RG transformation. The full line defines the critical subspace. Initial Hamiltonians in this subspace converge to the fixed point. An initial sub-critical Hamiltonian, 1, is, by subsequent transformations mapped on the Hamiltonians 2, 3, 4, etc., converging to the origin. An initial supercritical Hamiltonian, 1', is, mapped on the Hamiltonians 2', 3', 4', etc., moving off to infinity. The dashed line represents the locus of the laboratory system, with fixed *physical* coupling constants, as the temperature changed.

origin is a *fixed point* for the RG transformation, with *all* the supercritical Hamiltonians belong to the *basin of attraction* associated with that fixed point. These considerations can be repeated for supercritical Hamiltonians which end at some of the fixed points at the infinity.

The *critical* Hamiltonian stays at the critical line, so this line is an *invariant space* for RG transformation. Following K. Wilson, let us take for granted that under repeated mapping the critical Hamiltonian will converge to a *fixed point Hamiltonian*, $\mathcal{H}^*$. However, one can imagine several fixed points with its own basin of attraction. This picture is usually represented as a RG flow diagram, Fig. 12.12 The consideration above can be generalized for the case of multi-dimensional parameter space which we shall not do here.

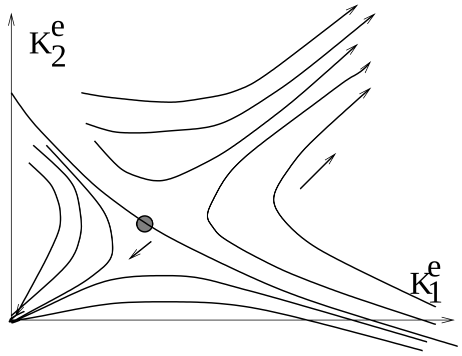

Figure 12.12: RG flow lines.

As we shall see, the critical exponents are determined by the properties of RG transformation near fixed points. That gives a new understanding of the concept *universality class*.

Fixed point properties

The definition of fixed point is

$$\mathbf{K}^* = \mathbf{K}'(\mathbf{K}^*) .$$

Near the fixed point one can linearize the mapping as

$$\delta K'_\alpha = \sum_\beta T_{\alpha\beta}\, \delta K_\beta , \quad \delta\mathbf{K} = \mathbf{K} - \mathbf{K}^* .$$

Let us assume that eigenvalues of the matrix $\hat{T}$ are real and positive. Since $\hat{T}$ is asymmetric one must discriminate between left and right eigenvectors. The left eigenvector belonging to the eigenvalue λ^i is defined as

$$\vec{\Phi}^i \cdot \hat{T} = \lambda_i \vec{\Phi}^i .$$

One can introduce the *scaling field* as

$$u_i = \vec{\Phi}^i \cdot \delta\mathbf{K} .$$

Then the rescaled field will be

$$u_i' = \vec{\Phi}^i \cdot \delta\mathbf{K}' = \vec{\Phi}^i \cdot \hat{T} \cdot \delta\mathbf{K} = \lambda_i \vec{\Phi}^i \cdot \delta\mathbf{K} = \lambda_i u_i .$$

Then one can assume that n repeated RG transforms with rescaling factor l are equivalent to a single transform with a scaling factor l^n,

$$[\lambda_i(l)]^n = \lambda_i(l^n) .$$

The solution of this equation has the form

$$\lambda_i(l) = l^{y_i}$$

where y_i is the critical exponent.

Usually 3 classes of the eigenvalues are discussed.

- *Irrelevant*, with $\lambda_i < 1$, i. e. $y_i < 0$. The corresponding scaling fields decay exponentially under RG transformations.
- *Relevant*, with $\lambda_i > 1$, i. e. $y_i > 0$. They grow exponentially and remove the system from the critical point.
- *Marginal*, with $\lambda_i = 1$, i. e. $y_i = 0$.

Connection to Physics

Scaling fields are convenient variable to analyze the equation

$$\mathcal{G}(\mathbf{u}) = \frac{1}{l^d} \mathcal{G}(\mathbf{u}') .$$

For magnetic system there are only 2 relevant fields, $u_\tau \to \tau$ and $u_h \to h$. Thus we have

$$\mathcal{G}(\tau, h) = \frac{1}{l^d} \mathcal{G}(\tau', h') = \frac{1}{l^d} \mathcal{G}(l^{y_\tau}\tau, l^{y_h} h) .$$

This is nothing else then the Widom's homogeneity postulate with replacement

$$l^d \to \lambda, \quad y_\tau/d \to a, \quad y_h/d \to b .$$

Summary

At present time, the RG theory is only an appealing picture, a seemingly fruitful way to think about critical phenomena. The precise conditions must be met to make the transform acceptable are less than clear. The main requirement is to leave long range fluctuations essentially untouched.

According to K. Wilson, “There is no RG cookbook!”.

Chapter 13

Transport phenomena

13.1 Classical transport

Boltzmann equation

Let us consider electrons in a solid state as an example. We shall concentrate on the transport under influence of electric and magnetic field.

According to classical physics, one can specify the coordinate **r** and the momentum **p** of the particle. Thus one can calculate the non-equilibrium distribution function, $f_{\mathbf{p}}(\mathbf{r},t)$, the current density being

$$\mathbf{j}(\mathbf{r},t) = e\int (d\mathbf{p})\, \mathbf{v}\, f_{\mathbf{p}}(\mathbf{r},t)\,.$$

Here we denote $(d\mathbf{p}) \equiv d^d p/(2\pi\hbar)^d$. The distribution function is defined by the Boltzmann equation

$$\begin{aligned}
\frac{d f_{\mathbf{p}}(\mathbf{r},t)}{dt} &\equiv \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{r}}\frac{d\mathbf{r}}{dt} + \frac{\partial f}{\partial \mathbf{p}}\frac{d\mathbf{p}}{dt} + I_{\text{coll}} \\
&= \frac{\partial f}{\partial t} + \mathbf{v}\frac{\partial f}{\partial \mathbf{r}} + \mathbf{F}\frac{\partial f}{\partial \mathbf{p}} + I_{\text{coll}} = 0
\end{aligned}$$

Here

$$\mathbf{F} = e\left(\mathbf{E} + \frac{1}{c}[\mathbf{v}\times\mathbf{H}]\right)$$

is the force acting upon the electrons, $\mathbf{v} \equiv \partial\varepsilon_{\mathbf{p}}/\partial\mathbf{p}$ is the (group) electron velocity, while I_{coll} is the *collision operator*. It expresses the changes in the state due to quantum collisions and can be expressed through the the transition probability W_{if} between the initial state (i) and the final state (f),

$$I_{\text{coll}}(f_\alpha) = \sum_{\alpha'}\left[W_{\alpha\alpha'} f_\alpha(1-f_{\alpha'}) - W_{\alpha'\alpha} f_{\alpha'}(1-f_\alpha)\right]\,.$$

Here α denotes the electronic state (in our case $\alpha \equiv \{\mathbf{p},s\}$).

In a stationary situation and in the absence of external fields the solution of the Boltzmann equation is the Fermi function. Substituting it into the collision operator we obtain a very important relation between the probabilities of the direct and reverse processes,

$$W_{\alpha\alpha'}e^{-\varepsilon_{\alpha'}/kT} = W_{\alpha'\alpha}e^{-\varepsilon_\alpha/kT}\,.$$

Drude formula

The simplest problem is to calculate the linear response to a small stationary electric field, **E**. Since we assume E to be small, it is reasonable to search solution as

$$f_{\mathbf{p}}(\mathbf{r}) = f_0(\varepsilon_{\mathbf{p}}) + f^{(1)}\,, \quad |f^{(1)}| \ll f_0\,.$$

Since $I_{\text{coll}}(f_0) = 0$ we come to the integral equation

$$I_{\text{coll}}(f^{(1)}) = -e\mathbf{E}\frac{\partial f_0(\varepsilon_{\mathbf{p}})}{\partial \mathbf{p}} = e\mathbf{E}\mathbf{v}\left(-\frac{d f_0(\varepsilon_{\mathbf{p}})}{d\varepsilon_{\mathbf{p}}}\right)\,.$$

The linearized collision operator has the simplest form for elastic processes when $W_{\mathbf{p}\mathbf{p}'} = W_{\mathbf{p}'\mathbf{p}}$. Then,

$$I_{\text{coll}}(f^{(1)}) = \sum_{\mathbf{p}'} W_{\mathbf{p}\mathbf{p}'}\left(f^{(1)}_{\mathbf{p}} - f^{(1)}_{\mathbf{p}'}\right)\,.$$

Since we are interested in the function odd in **p** (to create a current) and it must be a scalar it is natural to assume that

$$f^{(1)}_{\mathbf{p}} \propto (\mathbf{p}\cdot\nu)\,, \quad \nu \equiv \mathbf{E}/E\,.$$

Under such an assumption

$$I_{\text{coll}}(f^{(1)}) = \frac{f^{(1)}_{\mathbf{p}}}{\tau_{\text{tr}}}\,, \quad \frac{1}{\tau_{\text{tr}}} = \sum_{\mathbf{p}'} W_{\mathbf{p}\mathbf{p}'}\left(\frac{p_\nu - p'_\nu}{p_\nu}\right)\,.$$

The quantity τ_{tr} is called the transport relaxation time. If the material is isotropic then W is dependent only on the scattering angle θ between **p** and **p**$'$. From the simple geometrical construction shown in Fig. 13.1 we observe that

$$p'_\nu = p'\cos\theta\cos\phi\,, \quad p_\nu = p\cos\phi\,.$$

Consequently, we can express the relaxation time in the form

$$\frac{1}{\tau_{\text{tr}}} = g(\varepsilon_{\mathbf{p}})\frac{1}{2}\int_0^\pi d\theta\,\sin\theta\,(1-\cos\theta)\,W(\theta)\,.$$

Using the expression for the relaxation time we can write

$$f^{(1)}_{\mathbf{p}} = \tau_{\text{tr}}e(\mathbf{E}\cdot\mathbf{v})\left(-\frac{d f_0(\varepsilon_{\mathbf{p}})}{d\varepsilon_{\mathbf{p}}}\right) \quad \to \quad \mathbf{j} = e^2\int(d\mathbf{p})\,\mathbf{v}(\mathbf{E}\cdot\mathbf{v})\,\tau_{\text{tr}}(\varepsilon_{\mathbf{p}})\left(-\frac{d f_0(\varepsilon_{\mathbf{p}})}{d\varepsilon_{\mathbf{p}}}\right)\,.$$

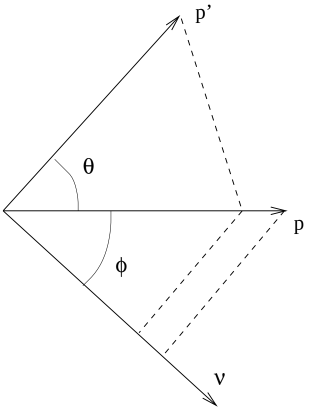

Figure 13.1: On the calculation of the transport relaxation time.

As a result, we arrive at the Ohm's law, $\mathbf{j} = \sigma \mathbf{E}$ with

$$\sigma = \int_0^\infty d\varepsilon\, g(\varepsilon) \langle v_{\mathbf{E}}^2 \tau_{\mathrm{tr}} \rangle_\varepsilon \left(-\frac{d f_0(\varepsilon)}{d\varepsilon} \right) . \tag{13.1}$$

Here $v_{\mathbf{E}}$ means the projection of the velocity on the direction of the electric field, while $\langle \ldots \rangle_\varepsilon$ means the average over the surface of a constant energy ε,

$$\langle A(\mathbf{p}) \rangle_\varepsilon \equiv \frac{\int (d\mathbf{p}) A(\mathbf{p}) \delta(\varepsilon - \varepsilon_{\mathbf{p}})}{\int (d\mathbf{p}) \delta(\varepsilon - \varepsilon_{\mathbf{p}})} .$$

The quantity

$$D = \langle v_{\mathbf{E}}^2 \tau_{\mathrm{tr}} \rangle_\varepsilon = \frac{1}{d} v^2 \tau_{\mathrm{tr}} = \frac{1}{d} v \ell$$

has an explicit physical meaning. This is just the diffusion constant for the electrons with a given energy ε. Indeed, for low temperatures we get

$$\sigma = e^2 g(\varepsilon_F) D(\varepsilon_F) . \tag{13.2}$$

On the other hand, the phenomenological expression for the current density in the presence of the density gradient reads as

$$\mathbf{j} = \sigma \mathbf{E} - eD \nabla n ,$$

where D is th diffusion constant. In the equilibrium the current is zero and

$$eD \nabla n = eD g(\varepsilon_F) \nabla \zeta = -\sigma \nabla \varphi .$$

At the same time, the electro-chemical potential $\zeta + \varphi/e$, must be constant and $\nabla \varphi = -e \nabla \zeta$. Thus we identify the quantity with the diffusion constant. Eq. (13.2) is known as the Einstein relation.

13.2 Ballistic transport

Landauer formula

We start this chapter by a description of a very powerful method in physics of small systems - so-called *Landauer approach*.

The main principle of this approach is the assumption that the system in question is coupled to large reservoirs where all inelastic processes take place. Consequently, the transport through the systems can be formulated as a quantum mechanical *scattering problem*. Thus one can reduce the non-equilibrium transport problem to a quantum mechanical one.

Another important assumption is that the system is connected to reservoirs by *ideal quantum wires* which behave as waveguides for the electron waves. We start our analysis from the discussion of the properties of an ideal quantum wire.

Ideal quantum wire

Consider 2 large reservoirs of electron gas reservoirs having the difference δn in the electron density and separated by a pure narrow channel. For small δn one can assume that there is a difference in a chemical potential, $\delta\mu = \delta n/g(\varepsilon_F)$. In the following we shall use the Fermi level of non-biased system as the origin for the chemical potentials. So the difference between the chemical potential in α-th reservoir will be denoted as μ_α.

If the channel is long and uniform, then the total current carried by the state characterized by a transverse mode n and a given direction of spin which propagates without scattering is

$$J_n = e\int \frac{dk_z}{2\pi\hbar}\frac{\partial\varepsilon_n(k_z)}{\partial k_z} = \frac{2}{2\pi\hbar}\int_{\varepsilon_F+\mu_\alpha}^{\varepsilon_F+\mu_\beta} d\varepsilon \frac{\partial\varepsilon_n(k_z)/\partial k_z}{|\partial\varepsilon_n(k_z)/\partial k_z|} = \frac{2}{h}\delta\mu .$$

If we take into account electron spin and N transverse modes are open, then the conductance is given by the expression $G = \frac{2e^2}{h}N$.

We come to a very important conclusion: an *ideal* quantum wire has *finite* resistance $h/2e^2N$ which is independent of the length of the wire.

As we have seen, even an ideal quantum wire has a finite resistance. That means a finite heat generation even in the absence of any inelastic processes inside the wire. Below we will discuss the physical picture of heat release by a current-carrying nanostructure.

First of all let us specify what heat release is. It will be convenient to consider an isolated system. Therefore we will have in mind the following physical situation. There is a capacitor which is discharged through the conductor of interest. The product RC of the whole system, R and C being the resistance and capacitance respectively, is much bigger than any relaxation time characterizing the electron or phonon system of the conductor. This means that for all the practical purposes the conduction process can be looked upon as a stationary one. The total energy of the system, $\mathcal{U}$, is conserved, while its total entropy, $\hat{\mathcal{S}}$, is growing. The rate of heat generation is expressed through $T\partial\hat{\mathcal{S}}/\partial t$, where T is the temperature, i.e. through the applied voltage and characteristics of the nanostructure itself. This means that the result is independent of the assumption that the considered system is isolated, which is made only for the sake of derivation. This thermodynamically defined heat is generated in the classical reservoirs over the length having a physical meaning of the electron mean free path. That is the same mean free path that enters the Drude formula, which determines the conductivity of the reservoirs themselves, the amount of heat generated per second in both reservoirs being the same.

It is interesting to indicate that even purely elastic collisions can result in a heat generation although they of course cannot establish full equilibrium. This has a clear physical meaning. The amount of order in the electron distribution resulting in electric current can bring about mechanical work. For instance, one can let the current flow through a coil, and a magnetic rod can be drawn into the coil. In such a way the electrons transferring the current can execute a work on the rod. As a result of scattering, the amount of order in the electrons' distribution diminishes, and this means dissipation of mechanical energy into the heat. It has been shown that the heat release is symmetric in both reservoirs even if the scatterers in the system are asymmetric.

All the above considerations do not mean that the collisions that give the main contribution to the heat release, also establish *full equilibrium*. What equilibrium needs is inelastic collisions which transfer the energy of electrons taking part in charge transfer to other degrees of freedom, such as to other electrons and phonons. In particular, a *local equilibrium electron distribution* is established over the length scale determined by *electron-electron* interaction. Such a distribution can be characterized by a local electro-chemical potential and sometimes an electron temperature. The latter can in principle be measured by optical methods. On the other hand, the *equilibrium with respect to the lattice* is established at the scales of *electron-phonon* and *phonon-phonon* mean free paths. Only over those distances from the channel one can treat the results in terms of the true local temperature.

Resistance of a quantum resistor

Consider a system shown in Fig. 13.2 consisting of a barrier connected to reservoirs by ideal quantum wires. If there is some reflection only a part of the current is transmitted. In this case one can introduce the transmission probability of the mode n, T_n, to obtain (including spin degeneracy)

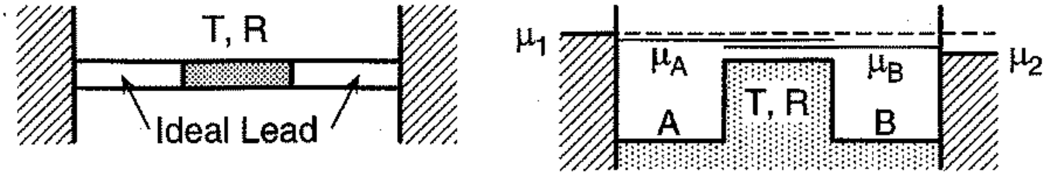

Figure 13.2: On the resistance of a quantum resistor.

$$J = \frac{2}{h}\delta\mu \sum_{n=1}^{N} T_n \,.$$

As a result,

$$G = \frac{2e^2}{h}\sum_{n=1}^{N} T_n = \frac{2e^2}{h}\mathrm{Tr}\,\mathbf{t}\mathbf{t}^\dagger \,. \tag{13.3}$$

Here **t** is the matrix of *scattering amplitudes* while the expression is called *two-terminal* Landauer formula.

This very important and looking simple formula was confusing during a long period. Indeed, this is the conductance which is measured between two reservoirs. Having in mind that the resistance of the connecting ideal wires (per one conducting mode) is $h/2e^2$ we can ascribe to the scattering region the resistance

$$\frac{h}{2e^2}\left[\frac{1}{T}-1\right]=\frac{h}{2e^2}\frac{R}{T},$$

where R is the reflection coefficient. Consequently, in the original formulation the quantum resistance was described as

$$G=\frac{2e^2}{h}\sum_{n=1}^{N}\frac{T_n}{1-T_n}. \tag{13.4}$$

However, the quantity which is usually measured is given by Eq. (13.3).

Now we derive the Landauer formula for finite-temperature and so-called multichannel case when the leads have several transverse modes. Consider ideal wires which lead to a general elastic scattering system. Let each lead has the cross section A and have $N_\perp$ transverse channels characterized by wave vectors k_i so that,

$$E_i+\frac{\hbar^2k_i^2}{2m}=E_F.$$

The incoming channels are fed from the electron baths with the same temperature and chemical potentials $\mu_1,\mu_2,\ldots$. The outgoing channels are fed up to *thermal equilibrium population*. We shall assume that the particles are absorbed in the outgoing baths. The sources are assumed to be *incoherent*, the differences $\mu_1-\mu_2$ are also assume small to yield linear transport. We introduce the scattering amplitudes t_{ij} for the transmission from jth incoming to ith outgoing channel. Reflection amplitudes r_{ij} are introduces in a similar way for reflection into the ith incoming channel. If we replace the incoming and outgoing channels, we denote the proper amplitudes by primes. In this way it is convenient to introduce $2N_\perp\times 2N_\perp$ scattering matrix as

$$S=\begin{pmatrix} r & t' \\ t & r' \end{pmatrix}.$$

From the current conservation we must require unitarity while from time reversal symmetry $S=\tilde{S}$. Thus we have also $SS^\star=I$ where star stays for complex conjugation while tilde for transposition. In a magnetic field the Onsager relation requires $S(H)=\tilde{S}(-H)$.

It one defines the total transmission and reflection into ith channel as

$$T_i=\sum_j|t_{ij}|^2, \quad R_i=\sum_j|r_{ij}|^2.$$

then from unitarity condition we get

$$\sum_i T_i=\sum_i(1-R_i).$$

Since the densities of states in each channel are 1D like, $g_i(E) = (\pi\hbar v_i)^{-1}$ we write the current through outgoing channels as

$$\begin{aligned} I &= \frac{e}{\pi\hbar}\sum_i \int dE\, \left[f_1(E)T_i(E) + f_2(E)R'(E) - f_2(E)\right] \\ &= \frac{(\mu_1-\mu_2)e}{\pi\hbar}\int dE\, \left(-\frac{\partial f}{\partial E}\right)\sum_i T_i(E)\,. \end{aligned}$$

Thus the conductance becomes

$$G = \frac{2e^2}{h}\int dE\, \left(-\frac{\partial f}{\partial E}\right) \mathrm{Tr}\,\mathbf{t}\mathbf{t}^\dagger\,.$$

This is the *two-terminal* conductance measured between the *outside* reservoirs which includes contact resistances.

Point ballistic contact

The most clean system is the so-called quantum point contact (QPC) - short and narrow constrictions in 2d electron gas. A sketch of QPC is shown in Fig. 13.3 The conductance of QPC is quantized in the units of $2e^2/h$. The quantization is not that accurate as in quantum Hall effect (about 1%) because of non-unit transparencies T_n and finite temperature. It is interesting to compare quantum and classical behavior of QPC. In a classical picture one can write

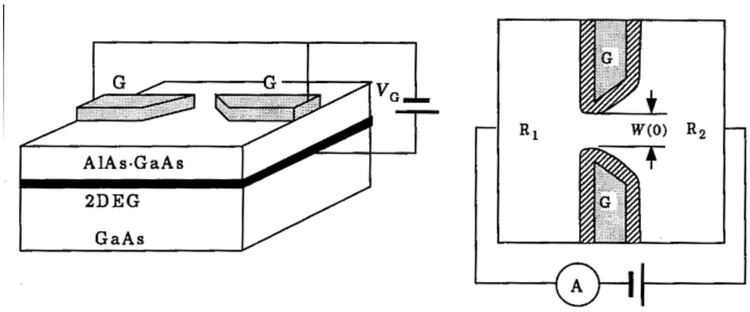

Figure 13.3: A sketch of a QPC formed by split gates.

$$J = W(\delta n)v_F \int_{-\pi/2}^{\pi/2} \frac{d\alpha}{2\pi}\cos\alpha = \frac{1}{\pi}Wv_F(\delta n)\,.$$

Thus the “diffusion constant” is

$$D_{eff} = \frac{J}{\delta n} = \frac{1}{\pi}Wv_F \quad \rightarrow \quad G = e^2 g(\varepsilon_F) D_{eff} = \frac{2e^2}{h}\frac{k_F W}{\pi}\,.$$

Note that the integer part of the quantity $k_F W/\pi$ is just the number of occupied modes according to quantum mechanics.

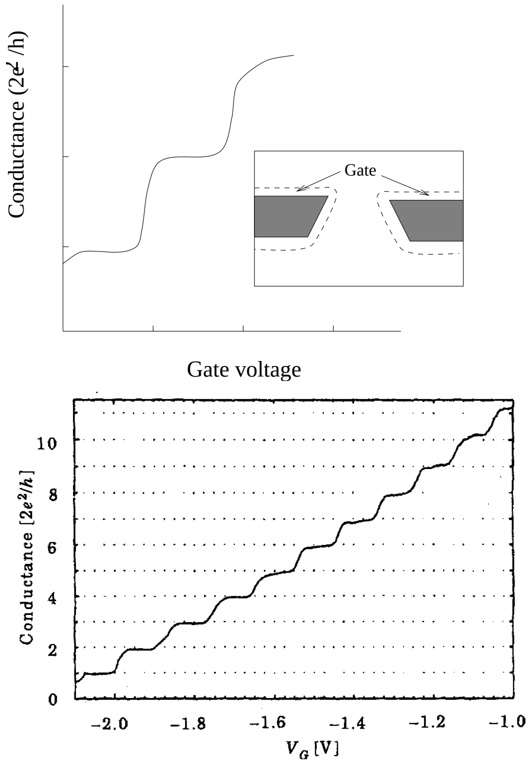

Figure 13.4: Quantization of conductance of a point contact: Schematic picture (left) and experimental result (right).

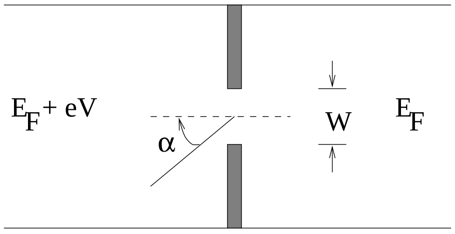

Figure 13.5: On the classical conductance of a point contact.

Appendix A

Notations

Microscopic variables

$q \to (q_1, q_2, \ldots, q_f)$ - generalized coordinates,
$p \to (p_1, p_2, \ldots, p_f)$ - conjugated momenta,
$\mathcal{H}(q,p)$ - classical Hamiltonian,
$\hat{\mathcal{H}}$ - quantum Hamiltonian,
$d\Gamma \equiv \prod_{j=1}^{f} dq_j\, dp_j$ - element of phase space,
m - mass of a molecule,
$\Gamma(E, \ldots)$ - number of states with the energy $\leq E$,
$D(E, \ldots) = \partial\Gamma/\partial E = \mathrm{Tr}\,\delta\left(E - \hat{\mathcal{H}}\right) = \frac{1}{V}\int d\Gamma\,\delta\left[E - \mathcal{H}(p,q)\right]$ - state density,
$\mathbf{S}$ - spin,
c - velocity of light,
g_i - degeneracy of i-th state.

Macroscopic variables

V - volume of the box which contains the system,
P - pressure,
T - absolute temperature (we use also energy units),
$\beta = 1/T$ - inverse temperature,
E - total energy,
U - internal energy, S - entropy,
S_t - total entropy for the body and surrounding medium,
F - (Helmholtz) free energy,
H - thermal function, enthalpy,
G - Gibbs free energy, thermodynamic potential,
Ω - Landau free energy.

Electric and Magnetic Variables

B - magnetic field,
e - electronic charge,
$\mu = e/2mc$ - Bohr magneton.

Constants

$k = 1.38 \times 10^{-16}$ erg/K - Boltzmann constant,
h - Planck constant, $\hbar \equiv h/2\pi$,

Useful mathematical formulas

The Stirling's formula:

$$\log n! \approx n \log n - n + 1 \quad \text{at} \quad n \gg 1 .$$

or more accurate

$$n! \approx \sqrt{2\pi \left(n + \tfrac{1}{6}\right)} n^n e^{-n} .$$

Thermodynamic relations

Thermodynamic potentials

Thermodynamic potential	Notation	Independent variables	Differential
Internal energy	U	S, V, N	$T\,dS - P\,dV + \mu\,dN$
Heat function (enthalpy)	H	S, P, N	$T\,dS + V\,dP + \mu\,dN$
Helmholtz free energy	F	T, V, N	$-S\,dT - P\,dV + \mu\,dN$
Gibbs free energy	G	T, P, N	$-S\,dT + V\,dP + \mu\,dN$
Landau free energy	Ω	T, V, μ	$-S\,dT - P\,dV - N\,d\mu$

Table A.1: Thermodynamic potentials (summary)

Basic distributions

The Gibbs distribution

$$w_n = \frac{1}{Z} e^{-\beta E_n} , \qquad Z = \sum_s e^{-\beta E_s} .$$

The Maxwell-Boltzmann distribution

$$\begin{aligned}
f(p,q) &= f_p(p)\cdot f_q(q)\,; \\
f_p(p) &= (2\beta\pi m)^{-3/2}\, e^{-\beta(p_x^2+p_y^2+p_z^2)/2m}\,, \quad \beta\equiv 1/T\,; \\
f_q(q) &= \left[\int dq\, e^{-\beta V(q)}\right]^{-1} e^{-\beta V(q)}\,.
\end{aligned}$$

The Fermi-Dirac distribution

$$f_k = \left[e^{\beta(\varepsilon_k-\mu)}+1\right]^{-1}.$$

The Bose-Einstein distribution

$$f_k = \left[e^{\beta(\varepsilon_k-\mu)}-1\right]^{-1}.$$

The chemical potential μ is determined by the normalization

$$V\sum_k \frac{g_k}{e^{\beta(\varepsilon_k-\mu)}\pm 1} = N\,.$$

Relations between thermodynamic and statistical quantities

$$\begin{aligned}
Z &= \sum_n e^{-\beta E_n}\,; \\
F &= -T\,\log Z\,; \\
\Omega &= -T\log\sum_N\left(e^{\beta\mu N}\sum_s e^{\beta E_{sN}}\right).
\end{aligned}$$

Appendix B

On the calculation of derivatives from the equation of state

Important relations arise from the properties of partial differentiations. Consider 3 quantities, X,Y,Z, related by the equation of state $\mathcal{K}(X,Y,Z) = \text{const}$. Now let us consider X,Y as independent variables, while $Z = Z(X,Y)$. We have

$$\left(\frac{\partial Z}{\partial X}\right)_Y dX + \left(\frac{\partial Z}{\partial Y}\right)_X dY - dZ = 0\,. \tag{B.1}$$

If Y,Z are taken as independent variables, then

$$-dX + \left(\frac{\partial X}{\partial Y}\right)_Z dY + \left(\frac{\partial X}{\partial Z}\right)_Y dZ = 0 \tag{B.2}$$

Now we multiply Eq. B.1 by $\left(\frac{\partial X}{\partial Y}\right)_Z$ and Eq. B.2 by $\left(\frac{\partial Z}{\partial Y}\right)_X$ and subtract the results. We obtain

$$\left[\left(\frac{\partial Z}{\partial X}\right)_Y \left(\frac{\partial X}{\partial Y}\right)_Z + \left(\frac{\partial Z}{\partial Y}\right)_X\right] dX + \left[-\left(\frac{\partial X}{\partial Y}\right)_Z - \left(\frac{\partial X}{\partial Z}\right)_Y \left(\frac{\partial Z}{\partial Y}\right)_X\right] dZ = 0\,.$$

Since dX and dY are *independent*, this equation is compatible if

$$\begin{aligned}
\left(\frac{\partial Z}{\partial X}\right)_Y \left(\frac{\partial X}{\partial Y}\right)_Z + \left(\frac{\partial Z}{\partial Y}\right)_X &= 0 \\
\left(\frac{\partial X}{\partial Y}\right)_Z + \left(\frac{\partial X}{\partial Z}\right)_Y \left(\frac{\partial Z}{\partial Y}\right)_X &= 0\,,
\end{aligned}$$

or

$$\left(\frac{\partial X}{\partial Y}\right)_Z \left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y = -1\,, \tag{B.3}$$

$$\left(\frac{\partial X}{\partial Y}\right)_Z = \left(\frac{\partial Y}{\partial X}\right)_Z^{-1}\,. \tag{B.4}$$

General scheme for transformation: Consider any quantity $\mathcal{F}(X,Y)$, the differential of which can be expressed as

$$d\mathcal{F} = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_Y dX + \left(\frac{\partial \mathcal{F}}{\partial Y}\right)_X dX\,.$$

Then we divide it by dZ and assume $Y = \text{const},\ (dY = 0)$. We get

$$\left(\frac{\partial \mathcal{F}}{\partial Z}\right)_Y = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_Y \left(\frac{\partial X}{\partial Z}\right)_Y . \tag{B.5}$$

Another important relation arises if one divides the expression for $d\mathcal{F}$ by xX,

$$\left(\frac{\partial \mathcal{F}}{\partial X}\right)_Z = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_Y + \left(\frac{\partial \mathcal{F}}{\partial Y}\right)_X \left(\frac{\partial Y}{\partial X}\right)_Z . \tag{B.6}$$

Equations (B.5), (B.6) together with Eqs. (B.3), (B.4) and the Maxwell relations are usually used for transformations and computation of derivatives from the equation of state.

Bibliography

[1] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd edition, Part 1 (Pergamon Press, Oxford - New York - Seoul - Tokyo), 1980.