THERMODYNAMICS AND STATISTICAL PHYSICS
LECTURES

Events important for students of Topics in Statistical Physics

- 4th December 2023 at 9.00 in room 1.03 - 1st colloquium
- 15th January 2024 at 9.00 in room 1.03 - 2nd colloquium
- 1st February 2024 at 9.00 in room 1.03 - 1st written exam

LITERATURE

- A.B. Pippard, The Elements of Classical Thermodynamics (number 1 of all times)
- H.A. Buchdahl, The Concepts of Classical Thermodynamics
- C.J. Adkins, Thermodynamics
- H. Callen, Thermodynamics
- E. Fermi, Thermodynamics
- R. Kubo, Thermodynamics (Problems with solutions)
- P. Landsberg, Thermodynamics and Statistical Mechanics
- R. Kubo, Statistical Mechanics An advanced Course (Problems with solutions)
- L.D. Landau & J.M. Lifshitz, Statistical Physics (Course of Theoretical Physics, vol. V)
- K. Huang, Statistical Mechanics

and in Polish

- J. Werle, Termodynamika
- K. Gumiński, Termodynamika
- K. Rejmer, Ciepło⟶Zimno, część pierwsza: Zasady, część druga: Zastosowania
- L.D. Landau, J.M. Lifszyc, Fizyka Statystyczna (t. V Kursu Fizyki Teoretycznej)
- K. Huang, Mechanika Statystyczna

Epigraph 1:
J'ai passionement aimée la Mediterrannée, sans doute parsque venue du Nord, comme tant d'autres, après tant d'autres...
Fernand Braudel, *La Mediterrannée et le Monde Mediterrannéen à l'Époque de Philippe II.*

Epigraph 2:
It may be objected by some that I have concentrated too much on the dry bones, and too little on the flesch that clothes them, but I would ask such critics to concede at least that the bones have an austere beauty of their own.
A. Brian Pippard, *The Elements of Classical Thermodynamics*

Epigraph 3:
There are only two true pillars of theoretical physics: quantum field theory and statistical mechanics; all the rest are dyrdymały.
I myself. Put it on my tomb.

LECTURE I (TMD)

Thermal physics or **statistical thermodynamics** in the large meaning of these terms encompasses - as Pippard puts it - all ways and methods of understanding and interpreting properties of matter in so far as they are influenced by changes of temperature. In this sense it is one of the major subdivision of physical science as such and it employs a variety of mathematical and experimental techniques, as well as all available information about microscopic constitution of matter to achieve its goal - the explanation of observed properties of matter at all temperatures and all conditions. This is, of course, what makes the subject so interesting as it encompasses essentially all physics (and besides it, also some parts of sciences like chemistry, biology and others) and is important to practically all its areas, from condensed matter to cosmology. Briefly, all physics unites in it. Thus, all general courses of classical mechanics, quantum mechanics, electrodynamics you have taken in this Department can be considered introductions to statistical thermodynamics.

This general goal has resulted in the quite widespread tendency (in fact not a modern one: look how the Statistical Physics volume - its first edition appeared in 1939! - of the famous Landau & Lifschitz theoretical physics course is organized) to present this subject mixing statistical physics with thermodynamics. I, however, prefer not to proceed in this way, and would like in these lectures to clearly separate what is called **classical thermodynamics**, to which a rather large first part will be devoted, from the **statistical physics** part. Perhaps this will make some of you unhappy - those who have already taken thermodynamics course run here by experimentalists and have hoped to have this awful thing "aus dem Kopf" once for ever - but I'm convinced (I'm not alone in this) this is the right way of presenting the subject. First of all, you have in fact only a very poor knowledge of what thermodynamics really is, and secondly, at this level one cannot go too far with the true statistical physics - all we are going to discuss will be the three statistical ensembles applied mainly to systems of nointeracting particles. Going further would require changing the format of this course to 2+3 at least[1] and would require a better theoretical background (starting from some level statistical mechanics becomes essentially the quantum field theory od many-body systems). But more importantly, a good working knowledge of thermodynamics is indispensable because when the statistical approach is applied to a physical system, the first goal is to recover its thermodynamics and this, once some basic functions (called thermodynamical potentials) characterizing this system are computed using statistical physics methods, is accomplished by applying to these functions the formalism of thermodynamics. So one has first to understand and master it.

Classical Thermodynamics (CTMD) originated in the XIX century from the study of machines like steam engines but it was soon recognized to be of such great generality that it applies to all phenomena and processes in which heat and temperature play important roles. In physics it provides general understanding of phenomena as different as thermal

---

[1] We have managed to convince the Dean to this idea and the parallel English version of this course runs already in the extended format.

radiation on one hand and low temperature properties of paramagnetic salts on the other. CTMD is a **phenomenological theory** - it does not delve into the microscopic structure of matter. Its goal is more modest: it is establishing **relations** between observed (measured) properties of bodies and substances thereby reducing the very large number of such properties to only a few; the other ones can be then treated as consequences of these few. For instance a gas exhibits different behaviours under various conditions, but given its equation of state in the form $f(T, V, p) = 0$ and one of its molar heats, say $c_p(T, p_0)$ over a range of temperatures (for a single fixed value p_0 of the pressure) it is possible to predict quantitatively how it will behave under various circumstances, e.g. how the temperature of the gas will change when it is passed through a throttle (a valve), how its (molar) volume will change with temperature, what is its molar heat c_v in the same range of temperatures for arbitrary molar volumes v, compute the work that can be extracted when the gas is expanded, etc. On the other hand, while taking (within the statistical approach) into account the atomic structure of matter it is possible to associate some particular kinds of macroscopic behaviour of physical systems - e.g. some anomalies in measured specific heats - with concrete microscopic features of these systems, this is not possible within pure thermodynamics.

CTMD deals only with equilibrium states of macroscopic physical systems and although it allows to make predictions concerning "processes" in which systems undergo some changes, following the real time evolution of the system's state is beyond its scope.[2] For this reason some would prefer the name thermostatic. Nevertheless, using essentially only idealizations like reversible processes, adiathermal isolations, etc. which are limiting cases of real situations, thermodynamics is able to formulate basic restrictions which apply to real processes and cannot be overcame.

Classical Thermodynamics essentially rests on only four (but in most cases three are sufficient) **laws** which I denote nTMDL, $n = 0, 1, 2, 3$. (The "law" means here something which cannot be derived from other rules.) These laws constitute a generalization of a great number of experimental observations. Continued application of the methods of classical thermodynamics based on these laws to all kinds of practical problems showed these laws give correct predictions in all cases. This is the empirical justification of these laws as having very large (practically unlimited) range of validity (they were never falsified - this is what the term "physical law" really means). The "heart" of thermodynamics is its second law which is associated with the mysterious quantity called entropy and to get acquainted with this notion we will discuss it in these lectures from several different points of view cosidering different formulations of 2TMDL.

Classical Thermodynamics does not attempt to explain the origin of these laws, that is, it does not inquire how they emerge as a consequence of more fundamental laws which govern the behaviour of microscopic constituents of matter - this is the goal of statistical

---

[2] There exists an extension of thermodynamics called **nonequlibrium thermodynamics** which allows to discuss on phenomenological ground also the behaviour of systems which are not in equilibrium but its range of validity is rather limited (in particular it is limited to systems which are "not too far" from equilibrium).

physics which we are going to treat in the second part of this lecture. But we will see that in fact they cannot be - at least to my taste - fully derived from the laws of mechanics (be it classical or quantum); they are rather replaced by other postulates (laws, but now ones formulated in more abstract terms) which again lead to correct predictions in most cases. However, in statistical physics the problem becomes more complicated because the practical application of its methods relies in most cases on approximations and/or on using simplified models of matter. Nevertheless, analyzing the behaviour of systems consisting of huge numbers of molecules (particles) it can be shown (as Pippard puts it: with a fair degree of rigor - enough to satisfy most physicists but few pure mathematicians) that those their general properties which can be treated as macroscopic and which follow only from statistical averaging, being to a large extent independent of the details of the microscopic dynamics, do indeed obey the laws of CTMD.

Of course statistical physics, apart from reproducing thermodynamics, providing inputs which otherwise (within purely thermodynamical treatment) would have to be laboriously reconstructed from experimental data and relating certain general types of observed thermodynamic behaviour (e.g. the already mentioned anomalies in specific heats) directly to concrete details of the microscopic structure of matter, allows also to ask and investigate theoretically questions which are entirely beyond the scope of classical thermodynamics. Here belong for example fluctuations, their spatial and temporal correlations, critical exponents (characterizing continuous phase transitions), etc. One can also formulate kinetic theories of various phenomena occurring in physical systems and study theoretically processes of approaching by these systems equilibrium states in various conditions.

One can therefore think, that classical thermodynamics, which was mainly developed when the microscopic constitution of matter was not yet fully investigated and its atomic basis was even questioned, is already “passé” and has become obsolete at least for those who want to inquire into the deepest mysteries of the physical world and to discover its most fundamental laws (laws of quantum gravity for example). However:

- Frequently, and almost as a rule in applied sciences, it is of primary interest to know relations between properties of substances rather than to know from which fundamental rules these properties follow.
- Thermodynamics greatly reduces the number of properties which have to be determined experimentally or have to be computed theoretically employing the statistical physics methods (as has been mentioned: statistical physics will give certain functions - thermodynamical potentials - from which all follows thermodynamically).
- As already said, employing statistical physics to real physical systems one is forced to make approximations or use simplified models; knowledge of thermodynamics allows us to clearly see which of the obtained properties are general and which are valid only in particular models.
- It is practically impossible to analyze with methods of statistical physics very complex systems (e.g. biological ones), whereas the simple rules of thermodynamics still

allow to make predictions concerning such systems and understand (even if only in very general terms) their behaviour.

- It sometimes turns out (unexpectedly) that thermodynamics does have something to do with the deepest fundamental mysteries - vide entropy of black holes (a brief account of it will be included in these lectures).
- Training of a physicist should involve not only learning fundamental theories, but also developing sensibility to the ways physical systems behave and for this thermodynamics provides a framework of very general ideas within which understanding of physical systems can be achieved.
- Last but not least, a less practical reason: the development of thermodynamical ideas has a formal elegance which is exceedingly satisfying aesthetically (Pippard again). It approaches the ideal of mathematical rigor (hence attempts to mathematicize it - we will not go this way!) closer than any other branch of natural science; furthermore, its historical development (fascinating as almost all history of science, but we have to omit it) and notions it introduces (entropy!) have become part of our culture - for all these reasons it should be an important part of education of a scientist (and definitely of any physicist).[3]

"Und somit fangen wir an." - as says Thomass Mann in the introduction ("Die Geshichte Hans Castorps die wir erzahlen wollen...") to his Zauberberg.

Two more quotations for the good start from two Arnolds:

- Thermodynamics is a funny thing: when you learn it for the first time, you don't understand it at all. On the second approach you have the impression of understanding it all, except for a few small details. At the third approach you already know that you don't understand it, but when you get accustomed to this, you cease to care about. **Arnold Sommerfeld**
- Every mathematician knows that understanding of an elementary course of thermodynamics is impossible. **Wladimir Igorievich Arnold**

I will, however, do my best to make thermodynamics as understandable for you as it can be. In general (this concerns also the statistical part of these lectures), I will try to at least give you a firm basis on which further study of thermodynamics and statistical physics can be pursued.

---

[3] One can add here that studying thermodynamics is an ideal place where to learn to clearly and in precise, well defined terms, express in words what is the meaning of mathematical steps taken and what is the logic behind them. This should help to tame the nowadays widespread tendency to produce only formulae without discussing the essence of physics considered, to use arbitrarily chosen terms to name things and to pay little - if at all - attention to the logic of presentations.

- **A system.** This portion of the Universe which is chosen for investigation. Here we will deal only with **macroscopic** systems (in line with the declaration that thermodynamics is not concerned with the microscopic constitution of matter). It will be convenient to distinguis two general kinds of systems:

- **Mechanical systems (in the generalized sense).** Systems of macroscopic material bodies moving under influence of their mutual interactions. These are captured by laws of classical theories like mechanics, elastomechanics or hydrodynamics, etc. Related to these theories are the familiar concepts of mechanical energy (kinetic and potential), work, forces etc. Mechanical energy of such systems is conserved (this is part of their definition): any work done on such a system by external forces is exactly equal to the change of the system's mechanical energy. If the initial and final state of a mechanical system are static (all velocities and accelerations of its parts vanish) the changes of its mechanical energy are fully characterized by changes of the system's **deformative coordinates** (variables) which are of a "geometrical" nature, like distances, volumes (if one includes also systems possessing electromagnetic properties, the system's magnetization and/or polarization should be also treated as deformation coordinates). Examples: two massive blocks on one another in the Earth gravitational field; vertical distance h between their centers is the deformation variable; if one mass is lifted (by an external force counterbalancing only the gravitation) the change of the system's energy is entirely determined by the change of h. Another one: an elastic rod of length L; if it is shortened or elongated by applying a force, its energy change depends solely on the change in L. Of course, mechanical systems are only an idealization.

- **Thermodynamic systems.** Mechanical energy of most systems is not conserved[4] (as experience shows). If a fluid contained in a vessel is at rest, its deformative coordinate being the volume V, its mechanical energy can be taken as zero. After stirring it, when the fluid is back at rest, its mechanical energy is still zero, yet some work has been done on it: to fully characterize the fluid's state yet one more variable, called **nondeformative**, is needed such as the fluid pressure p. In general, thermodynamic (noncompound) systems are ones the full characterization of static conditions of which requires a single nondeformation variable. A thermodynamic system can be:

- **Simple.** A single homogeneous body or a substance; only two parameters (coordinates), at least one nondeformative, are needed to fully characterize its equilibrium state - the most widely used example (the "working horse" of the elementary thermodynamics) is the fluid (a liquid or a gas - usually the perfect one) characterized by the variables V (deformative) and p (nondeformative). One can also consider magnetic substances characterized by the variables like $\mathbf{M}$ - its total magnetization

---

[4] In mechanics they are classified as dissipative systems.

(deformative) and $\boldsymbol{\mathcal{H}}_0$ - the strength of the applied external magnetic field (nondeformative) or dielectric materials - the variables are then $\mathbf{P}$ - its total polarization and $\boldsymbol{\mathcal{E}}$ (in both these cases when the volume effects can be neglected, e.g. at zero pressure).

- **Not simple.** It is also single homogeneous body or a homogeneous substance, but more than two parameters are needed (e.g. two deformative variables and a single nondeformative one); it will be seen later that the number of variables (parameters) needed to specify its equilibrium state is equal 1 plus the number $o-1$ of ways a quasistatic work can reversibly[5] be performed on the system. The deformative variables (in contrast to the constitutive ones - to be defined below - are "external" in the sense that their changes are directly related to interactions of the system with its surrounding. The condition that both, simple and not simple, thermodynamic systems require always only one nondeformative variable means that they cannot have internal adiathermal partitions.

- **Compound (non-homogeneous).** Some systems may consist of **several homogeneous parts** separated from one another by (or in contact with one another through) some walls of definite properties (to be defined below) alowing or inhibiting different kinds of interactions between them. They may be formed by several bodies or substances each of which is a simple or not simple (e.g. two containers with same gas or different gases of which one with magnetic properties). Sometimes also a single nonhomogeneous body (e.g. a gas in the gravitational field) must be mentally split into small parts which can be treated as homogeneous to make thermodynamics applicable to it.

- **Physico-chemical systems.** Systems as defined above (simple or not simple) are treated as "black boxes" - no variables are associated with their internal composition (even if some changes in their internal constitution may occur as their parameters vary) and the amount of matter in each of them is assumed to be constant. However, in some applications - in particular in chemistry - one may be interested in the internal composition which may be different in different equilibrium states as a result of reactions occuring within the system (matter can also be exchanged beetween different subsystems of compound thermodynamic systems). In such a case one needs to characterize states of the considered systems by specifying their internal constitution with the **constitutive coordinates**: a single n - the number of moles of the substance - in the case of chemically pure systems (e.g. pure water) or $n_1, \ldots, n_r$ in the case of homogeneous mixtures of r components[6] - e.g, a mixture of phenol - C_6H_5OH - and water). This allows then also to admit an exchange of matter between the system and its surrounding or between different parts of the

---

[5] This notion will be explained in due course.

[6] One can also use $n = \sum_{i=1}^{r} n_i$ - the total number of moles and the molar fractions $x_1, \ldots, x_r$, $x_i = n_i/n$, $\sum_{i=1}^{r} x_i = 1$.

same compound system. Some systems, e.g. the electromagnetic field, cannot be characterized by constitutive variable(s).

- **Composed of Phases.** Under certain conditions a given homogeneous system may split into several phases that is, into separate homogeneous parts (this means that splitting increases the number of parts of the system) with definite boundaries. This happens when in the system a phase transition of the first order occurs. A phase may be chemically pure (one material constituent) or be a mixture - e.g. the phenol-water mixture can split into phases with different concentrations in each. Discussing phases one necessarily ceases to treat the thermodynamic system as a black box. (One speaks also of different phases of a given system when a continuos transition occurs in it but in this case the phases do not form separate homogeneous parts.)

What is not included in the system constitutes its **surrounding**. More precisely, as the surrounding one takes into account only those parts of the rest of the Universe which may somehow influence the system under study (here the physical intuition becomes indispensable). In thermodynamical considerations the surrounding is usually modeled as consisting of sources of work, heat and matter, which may be exchanged with the system under study.

Boundaries separating the system from its surrounding or different parts of a compound system can be natural as in the case of the water droplet but frequently they are artificial or just mental constructions, and are generally called **walls** (or **partitions**, if they separate two parts of a compund system). They should be treated as constraints the studied system is subjected to. Walls play crucial roles in various thermodynamical reasonings. They can be of different character and can allow or inhibit different kinds of interactions of the system with its surrounding or between different parts of the same (compound) system. For instance, a rigid wall prevents performance of a mechanical work which requires changing the volume or the shape of the system, i.e. of the so-called volume work, on the system (gas, liquid, solid) enclosed by it. Walls can also allow or inhibit an exchange of matter (of all kinds of matter or of only one particular kind of it) between the system and its surrounding or between different parts of a compound system.

Of particular importance for many thermodynamical reasonings are the **adiathermal walls**.[7] Normally we would say such walls inhibit a heat transfer to or from the system (thermal interaction of the system with its surrounding), but since we (officially) don't know yet what the heat is, to explain this notion it is better to use the words of Pippard who says: "the walls of different vessels differ considerably in the ease with which influences from without may be transmitted to the system within. Water within a thin-walled glass flask may have its properties readily changed by holding the flask over a flame or putting it into a refrigerator; or the change brought about by the flame may be simulated

---

[7] I deliberately use this term - in place of the commonly used "adiabatic" - to have a clear opposition: diathermal and adiathermal.

(though not so easily) by directing an intense beam of radiation onto the flask. If, on the other hand, the water is contained within a double-walled vacuum flask with silvered walls (Dewar vessel), the effect of the flame or refrigerator or radiation may be reduced almost to nothing. (...) it is not a very daring extrapolation to imagine the existence of a vessel having perfectly isolating walls, so that the substance contained within it is totally unaffected by any external agency" (except the gravitational field). Another way of characterizing adiathermal walls is to say that a system in the adiathermal enclosure impermeable to matter can be disturbed only by mechanical means (which may but may also not be associated with changing its deformative coordinates; if these do not change, one speaks of an "isometric change"). Walls which do not have this property are called **diathermal**. Two systems contacted with one another through such a wall are said to be in **thermal contact**.

One should add that walls which are adiathermal in the above sense and rigid can still allow for some kinds of work to be done on the system, e.g. electrically (by passing a current through a resistant wire inserted into the system), or by stirring.

The set of variables characterizing a state of a thermodynamic system consist originally of deformative variables (like volume V in the case of simple fluids, magnetization $\mathbf{M}$) and at least one nondeformative (like pressure p or magnetic field strength $\boldsymbol{\mathcal{H}}_0$, etc.) all of which are also of "mechanical" nature (and are directly measurable). In the course of the development of the theory one defines new thermodynamical quantities like temperature T, chemical potential(s) μ, enthalpy H, entropy S and others, which can be used to characterize the system's state in place of the original ones. Some of the variables can be combined into pairs - the rule here is that the product of the variables forming a pair has the physical dimension of energy - and are said to be conjugated to one another. Such pairs are for instance p and V, $\boldsymbol{\mathcal{H}}_0$ and $\mathbf{M}$, T and S; in general the first variable of such a pair has the character of a (generalized) force and the second one (deformative) of a (generalized) displacement. All thermodynamic variables fall into two classes: **intensive** ones of essentially local character (like p, $\boldsymbol{\mathcal{H}}_0$, μ) and **extensive** ones characterizing the system as a whole (the system's total mass M, its volume V, internal energy U). If the considered system does have the property of **extensiveness** (most systems which will be considered do have it), then variables belonging to the latter class are proportional to the amount of matter in the system (or to its size). It is then convenient to operate with specific or molar quantities which are extensive quantities referred to a unit of mass or to one mole, respectively.[8] I denote molar quantities by a lower case character (e.g. c_v, c_p, u), and the trully specific ones, referred to a unit mass, by a tilde (e.g. $\tilde{c}_v$, $\tilde{u}$, etc.). A quantity the value of which depends uniquely on the (equilibrium) state of the system is called a **state function**.

Equilibrium

All classical thermodynamics rests on the following empirical fact: almost every system

---

[8] Sometimes, somewhat incorrectly, molar heat capacities, etc. are called "specific"; at any rate converting molar quantities into specific ones amounts only to multiplying by a constant (for a given substance) factor.

shielded by adiathermal walls and not subjected to any external influences tends towards and eventually (sooner or later) reaches a state in which no further change (of its macroscopically defined characteristics) is perceptible, no matter how long one observes it. This state is called the **thermodynamical equilibrium** state. In general it is only the state of equilibrium that can be characterized by only a small number of quantities (intensive or extensive). One also requires that there be no macroscopic flows in the system like a steady flux of heat, or an electric current passing through it. Water flowing down a tube is not strictly speaking a system in equilibrium - it is in the so called steady state; we will see, however, that to some extent TMD can be applied to it. Systems which cannot attain an equilibrium state in given conditions are excluded from thermodynamics considerations.

If the system is compound (has several parts or some heat or work sources are included in it), in approaching its equilibrium state changes in its individual parts usually occur - they can exchange heat between them[9] (e.g. a mercury in glass thermometer inserted in water changes but after a while it stabilizes), or perform works on one another but eventually attain equilibrium. If the external conditions, or walls through which various parts of a compound system communicate, are changed that is, the **constraints** to which the system is subjected are changed, a new equilibrium state is attained and one of the main roles of thermodynamics is to determine this new state (the Callen's point of view which we are going to discuss in due course).

Similarly as in mechanics one can contemplate different kinds of thermodynamical equilibria:

- stable mechanical (a ball resting in the global minimum of a potential) - a TMD analog is e.g. a pure gas at uniform temperature and pressure in a cylinder: upon a small disturbance (e.g. the pressure and density of a gas is made somewhat nonuniform by a short external perturbation) it reaches back the same equilibrium state.

- neutral mechanical (a ball on a flat table) - a mixture of water and its (saturated) vapour in a cylinder at special T and p - moving the piston causes a change in the proportions of water and vapour, but the system does not return to the previous state.

- metastable mechanical (a ball in a local minimum of a potential) - e.g. supercooled vapour or a mixture of hydrogen H_2 and oxygen O_2; they look as fully stable - can remain unchanged perceptibly for a very long time $\sim 10^{100}$ years!, and can even by subject to small disturbances (e.g. of pressure or temperature) - but the effects of a nucleation center or a spark show these systems are in fact not in equilibrium states. For many *practical* purposes, however, they can be, if not artificially perturbed, treated as being in full equilibrium.

---

[9] To avoid the an yet undefined term “heat” one should rather say “can interact through a diathermal wall”.

- unstable mechanical - no analogy.

In fact no unstable equilibrium exists in mechanics of real systems too: it is a purely mathematical concept, for it relies on many idealizations (pointlike material bodies, neglecting internal structures of bodies, neglecting minute external influences). Real physical systems can only be in stable or metastable equilibria, but the range of displacements which do not cause them to leave their metastable state may be so narrow that we classify them as unstable.

It will be seen that similarly to mechanics, also in thermodynamics equilibrium states of a system that is realized in specific conditions (specific constraints) minimize or maximize suitable functions called **thermodynamics potentials** (examples are entropy, Helmholtz free energy, Gibbs potential, etc).

As a matter of facts, equilibrium of thermodynamical systems is never truly static: microscopic examination of the behaviour of particles or molecules constituting macroscopic systems reveals they are always in a state of a continuous agitation (Brownian motions); for instance the local density of a fluid always **fluctuates** a little around its mean value. If one waits long enough one might have a chance - theoretically only, because the times involved are typically much much longer than the Universe's lifetime! - of observing a sizeable departures from the mean state of the system; e.g. a gas filling 1 ccm can spontaneously contract to half of this volume and then in less than 10^{-4} sec. revert to its average density, but this can occur - it can be estimated - once in $10^{10^{19}}$ years. Of course the numbers quoted here cannot be obtained in pure thermodynamics, one needs to delve into the microscopic dynamics, that is, go over to the statistical theory (to the kinetic theory of gases). Using this theory, in addition to being able to compute such numbers, one draws a lesson that such fluctuations should be treated as an inherent feature of every equilibrium state. In general however, for most purposes after a reasonably short time one can treat every macroscopic physical system as having attained an equilibrium state (corresponding to the conditions the system is subjected to). If the consequences of treating it as such are not corroborated by experiment, we must revise this assumption - evidently the system has not yet attained equilibrium or its fluctuations play a crucial role in the phenomenon we want to explain.

0TMDL and the (empirical) temperature

To introduce and discuss the concept of temperature, central to the entire TMD, it is convenient to concentrate first on an especially simple system - a homogeneous fluid (a liquid or a gas). Its simplicity derives from the fact that the shape of the container is irrelevant - deformations of its shape which do not cause any change of the volume do not require any work. (In contrast, the shape of a solid can only be changed by the application of a stress and its thermal properties are then usually also affected.)

We accept a fact of experience that every equilibrium state of a fixed mass of a fluid is completely specified by its volume V and its pressure p (we assume the fluid has no electric or magnetic properties or that external electromagnetic fields are absent altogether so that these properties do not play any role).

We can fix the volume V of the fluid kept in a cylinder and adjust its pressure p to any desired value by placing the fluid (contained within diathermal walls) in an oven or a refrigerator or by any other means. The **important point** is: whatever the process by which the given values of V and p have been reached, the final state is always the same - same colour, smell, sensation of warmth, thermal conductivity, viscosity etc. Every macroscopic property that can be measured as characterizing this system can be treated as a unique function of V and p.

Let us now take any two simple systems, e.g. two fluids for definiteness. If they are both isolated and reach equilibrium states separately, and are then brought into thermal contact through a diathermal rigid wall (so that they cannot perform volume works on one another; we assume also that no other works are being performed on these systems either), changes in general will be observed to occur in both of them until they reach, as a compound system, a new equilibrium state. We say they are then in **thermal equilibrium** with one another.

On the basis of our sensory experience we are tempted to say that this is so because initially their temperatures were not equal. But the word "temperature" has no content yet. What is for the moment important is that two systems may be separately in equilibrium but not in equilibrium with one another.

Consider now two masses of fluids, adjusted to have V_1 and p_1 and V_2 and p_2, respectively, each in equilibrium in itself. They in general will not be in equilibrium with one another unless V_1, p_1, V_2 and p_2 are appropriately correlated: if we fix V_1, p_1 and V_2 then p_2 has to be adjusted in order the two systems are in equilibrium with one another (so that no changes in their states are observed when they are brought into thermal contact). So, their equilibrium requires that a relation of the general form

$$F(V_1,p_1,V_2,p_2)=0\,, \tag{1}$$

holds. The form of the function F can be determined experimentally and depends, of course, on both systems. To introduce the concept of temperature one has to prove that this relation always takes the form

$$\phi_1(V_1,p_1)=\phi_2(V_2,p_2)\,, \tag{2}$$

with the function ϕ_1 being a property of only the first system and ϕ_2 of only the second one. This can be achieved by relying on

0TMDL

If of the three bodies A, B and C, the bodies A and B are each separately in thermal equilibrium with the body C, then A and B are also in thermal equilibrium with one another.

It is useful to have it stated somewhat differently:

Converse 0TMDL

If the three bodies A, B and C are pairwise in thermal contact through appropriate diathermal walls and the whole system is in equilibrium, then any two of these bodies taken separately are also in equilibrium with one another.

One can illustrate the stated 0TMDL with an example: let C be a mercury in glass thermometer in which a quantity of Hg is roughly at zero pressure (the thermometer tube is evacuated) so that h - the height of the column of Hg in the tube - fully specifies its state. If the height is the same when the thermometer is put in thermal contact first with the body A and then with the body B, then nothing will happen when A is brought into thermal contact with B.

So let's pass to the reasoning. We first give a physical argument and only then sketch the mathematical one (which is nice, so worth showing).

Let us take two masses of fluids: S - the standard one which is kept at fixed V_S and p_S, and T - the one which is being tested. Vary V_T and p_T in such a way as to maintain equilibrium between T and S. In the plane (V_T, p_T) this determines a curve which will be called the **isotherm**. The isotherm does not depend on the standard body, for if another body S' is taken which is in thermal equilibrium with S, by virtue of 0TMDL, the same isotherms of the system T will be obtained. In other words, the isoterms reflect the properties of the test body itself.

Changing now V_S and p_S we can produce many isotherms of the tested body T. One can now introduce a system (however arbitrary) of labeling these isotherms by numbers t. In this way we define a function (which need not even be analytic at this stage but to avoid complications we will assume it is such)

$$\phi_T(V_T, p_T) = t\,,$$

and we call t the **empirical temperature**. Once this is done, we can produce isotherms of other simple bodies taking the test one for the standard. But if a consistency is to be achieved, there is no more freedom in labeling isotherms of other bodies: with each such a body there must be associated a function $\phi_{\rm body}$ depending on the parameters fully characterizing equilibrium states of that body and this function must take values equal $\phi_T(V_T, p_T) = t$ if the body is in equilibrium with the test body at V_T and p_T. In this way one establishes the existence of another state function - the (empirical) temperature t. It should be also clear that the reasoning extends to nonsimple bodies (systems): if more than two parameters are needed to completely characterize equilibrium states of a such a body, say $X_1, \dots, X_{o-1}, y$ (X_i are its deformative coordinates and y is the nondeformative one), its isotherms are not curves, but hypersurfaces of dimension $o-1$ (or of codimension 1) determined by the equation

$$\phi_{\rm body}(X_1, \dots, X_{o-1}, y) = t\,. \tag{3}$$

A relation of this sort, more frequently written in the form $f(t, X_1, \dots, X_{o-1}, y) = 0$ is called the **equation of state** of the body (substance, system).

Now the mathematical reasoning. Consider three fluids A, B and C. If A and C are in equilibrium with one another, then

$$F_1(V_A, p_A, V_C, p_C) = 0\,, \qquad \text{so} \qquad p_C = f_1(V_C, V_A, p_A)\,.$$

Similarly, if B and C are in equilibrium with one another,

$$F_2(V_B,p_B,\,V_C,p_C)=0\,,\qquad\text{so}\qquad p_C=f_2(V_C,V_B,p_B)\,.$$

It then follows that

$$f_1(V_C,V_A,p_A)=f_2(V_C,V_B,p_B)\,.\tag{4}$$

But according to 0TMDL A and B are also in equilibrium with one another, so

$$F_3(V_A,p_A,\,V_B,p_B)=0\,,$$

and (4) must be equivalent to this relation. But since V_C does not enter F_3, it must drop out from the relation $f_1=f_2$. This is so if $f_1(V_C,V_A,p_A)=\phi_1(V_A,p_A)\psi(V_C)+\eta(V_C)$ and $f_2(V_C,V_B,p_B)=\phi_2(V_B,p_B)\psi(V_C)+\eta(V_C)$ with some universal functions $\psi(\cdot)$ and $\eta(\cdot)$. The relation $f_1(V_C,V_A,p_A)=f_2(V_C,V_B,p_B)$ is then equivalent to $\phi_1(V_A,p_A)=\phi_2(V_B,p_B)$.

That 0TMDL enforces dropping out of V_C, or rather that the three relations $F_i=0$ can be reduced to the form (2), can be shown more formally as follows. From $F_3(V_A,p_A,V_B,p_B)=0$ one can get $p_B=f_3(V_B,V_A,p_A)$ and write the equality (4) in the form

$$f_1(V_C,V_A,p_A)=f_2(V_C,V_B,\,f_3(V_B,V_A,p_A))\equiv p_C\,,$$

which clearly shows that the right hand side must be independent of the variable V_B. So one can fix it and forget it changing the notation to

$$\begin{aligned}&f_3(V_B,V_A,p_A)\equiv g_A(V_A,p_A)\,,\\&f_2(V_C,V_B,\,f_3(V_B,V_A,p_A))\equiv\tilde{f}_2(V_C,\,g_A(V_A,p_A))\,,\end{aligned}$$

so that now the relation $p_C=f_2(V_C,V_B,\,f_3(V_B,V_A,p_A))$, from which V_B, as argued, drops out, can be written as

$$p_C=\tilde{f}_2(V_C,\,g_A(V_A,p_A))\,.$$

Disentangling from it g_A, one can write it in the form

$$g_A(V_A,p_A)=g_C(V_C,p_C)\,.$$

In the analogous manner one arrives at two similar relations

$$h_B(V_B,p_B)=h_A(V_A,p_A)\,,\qquad\text{and}\qquad t_B(V_B,p_B)=t_C(V_C,p_C)\,.$$

But we don't know yet, whether e.g. the function $t_B(V_B,p_B)$ can be related to $h_B(V_B,p_B)$. One still has to show that these three relations between the six functions can be reduced to three relations involving only three functions each of which depends on only one pair of the variables V, p.

To show this one can extract $p_C = \psi(V_C, V_A, p_A)$ from $g_A(V_A, p_A) = g_C(V_C, p_C)$ and put it into $t_B(V_B, p_B) = t_C(V_C, p_C)$:

$$t_B(V_B, p_B) = t_C(V_C,\, \psi(V_C, V_A, p_A))\,.$$

This shows that if V_A, V_B, V_C and p_A are taken for independent variables (in other words we treat p_B as determined by these four - we know from the previous considerations that if the bodies A, B and C are to remain pairwise in equilibrium, only four variables out of six can be varied independently), the right hand side of the relation $t_B = t_C$ is independent of V_B, so differentiating both sides of this equality w.r.t. V_B one obtains

$$\frac{\partial t_B}{\partial V_B} + \frac{\partial t_B}{\partial p_B}\frac{\partial p_B}{\partial V_B} = 0\,.$$

In turn, differentiating w.r.t. V_B the relation $h_B(V_B, p_B) = h_A(V_A, p_A)$ (continuing to treat V_A, V_B, V_C i p_A as independent variables; the right hand side of this relation is from the outset independent of V_B) one gets

$$\frac{\partial h_B}{\partial V_B} + \frac{\partial h_B}{\partial p_B}\frac{\partial p_B}{\partial V_B} = 0\,.$$

Eliminating now from these two relations the derivative $\partial p_B/\partial V_B$ one finds that

$$\frac{\partial t_B}{\partial V_B}\frac{\partial h_B}{\partial p_B} - \frac{\partial t_B}{\partial p_B}\frac{\partial h_B}{\partial V_B} \equiv \frac{\partial(t_B, h_B)}{\partial(V_B, p_B)} = 0\,.$$

Vanishing identically (i.e. for any values of the four independent variables) of this Jacobian means in effect that the mapping $(V_B, p_B) \longrightarrow (t_B, h_B) \in \mathbb{R}^2$ is degenerated (it is of rank 1, instead of being of rank 2), or - saying it more accessibly to the audience - the image in $\mathbb{R}^2$ of this mapping (defined on $\mathbb{R}^2$) is a one-dimensional curve, and not a two-dimensional domain. This in turn means that the functions $h_B(V_B, p_B)$ and $t_B(V_B, p_B)$ are not independent: there must exist a relation $r(t_B, h_B) = 0$ which determines the mentioned curve and this relation can again be inverted to give $h_B = \chi(t_B)$ that is, $h_B(V_B, p_B) = \chi(t_B(V_B, p_B))$. This can now be exploited in the relation linking the functions h_A and h_B: if we define a new function $t_A \equiv \chi^{-1}(h_A)$, this relation takes the form

$$t_A(V_A, p_A) = t_B(V_B, p_B)\,.$$

And this, combined with the equality $t_B(V_B, p_B) = t_C(V_C, p_C)$ yields $t_A(V_A, p_A) = t_C(V_C, p_C)$. This means that the third equality, $g_A(V_A, p_A) = g_C(V_C, p_C)$, must be equivalent to this one, that is $g_A(V_A, p_A) = r(t_A(V_A, p_A)) = r(t_C(V_C, p_C)) = g_C(V_C, p_C)$.

Scales of temperature

It is because of the freedom in labeling the isotherms of the test body (in the first reasoning presented above) or because one can always use $\tilde{t}_A = f(t_A)$, $\tilde{t}_B = f(t_B)$ and

$\tilde{t}_C = f(t_C)$ (in the mathematical reasoning) that the quantity t (a state function) is called the empirical temperature. We will see that TMD itself provides a mean of defining the **absolute** or **thermodynamic** temperature (it will be denoted T) the only freedom in its definition being that of scale. The practical question therefore is how to relate various empirical temperatures defined by different thermometric bodies to the thermodynamic temperature, i.e. how to calibrate thermometers used in practical measurements with respect to that temperature.

Usually one chooses a thermometric body with suitable properties and labels its isotherms by changing x - one of the parameters characterizing its equilibrium states, while keeping the other parameters fixed. The empirical temperature t can be then taken to be related to x by $x = f(t)$ with $f(\cdot)$ an arbitrary monotonic function. A particularly simple function is $f(t) = at$. There are then two ways of fixing the proportionality constant a. Either one chooses two points and *fixes the number of units* of t between them or one *ascribes a concrete value* of t to one particular point.

For instance, one takes the Mercury in glass at pressure $p = 0$, sets $h_{\rm Hg} = a_{\rm Hg} t_{\rm Hg}$ ($h_{\rm Hg}$ being the height - above some conveniently chosen reference level - of the column of Mercury in an evacuated glass tube at $p = 0$) and determines $a_{\rm Hg}$ by requiring that there be 100 degrees between the melting point of ice and the boiling point of water, both at normal pressure 1.013×10^5 Pa (the famous 1013 HPa $\equiv$ 1 atm). However, if one takes another thermometric substance, e.g. the ethyl alcohol and defines $t_{\rm Alc}$ through $h_{\rm Alc} = a_{\rm Alc} t_{\rm Alc}$ in the analogous way, both thermometers can be made (by the appropriate choices of the reference levels of the heights $h_{\rm Hg}$ and $h_{\rm Alc}$) to yield $t_{\rm Hg} = t_{\rm Alc}$ at the two chosen reference points but will in general differ (even if not too much in practice) at all other points ($t_{\rm Hg} \neq t_{\rm Alc}$). This is because $t_{\rm Hg}$ and $t_{\rm Alc}$ are two different empirical temperatures.

Another choice of the thermometric substance is gases at very low pressures. This is a very convenient choice, because isotherms of gases under this condition are simple: $pV =$ const. This is the empirical Boyle-Marriotte law which is satisfied to a very good accuracy by real gases at sufficiently low pressures. It is therefore natural to set $pV = f(t)$, where t is the empirical temperature and f an arbitrary function. It turns out that if the empirical temperature scale $t_{\rm Hg}$ is fixed by a Mercury in glass thermometer, the function $f(\cdot)$ such that $pV = f(t_{\rm Hg})$ is nearly linear ($f(t_{\rm Hg}) = a\, t_{\rm Hg} + b$) over a wide range of temperatures. Therefore one sets $pV = nR\, t_{\rm gas}$ (n being the number of moles of the gas) and keeping e.g. the pressure fixed determines the value of the factor R in one of the two ways explained above. If the ice and the boiling water points at $p = 1$ atm are taken to calibrate the perfect gas scale and correspond respectively to the gas volumes V_1 and V_2 of a gas thermometer kept at the same (low) pressure, then $t_{\rm gas} = V \cdot 100/(V_2 - V_1)$. It should be noticed that in this definition of $t_{\rm gas}$ there is no freedom to fix a reference value (as it is possible with the Mercury in glass thermometers by choosing the reference height from which the height of the Mercury column is read). It then turns out that $t_{\rm gas}$ and $\tilde{t}_{\rm gas}$ obtained in this way using different gases (at sufficiently low pressures) are nearly the same not only at the chosen reference points but in a wide range of temperatures.

This is of course because the temperature $t_{\rm gas}$ defined in this way using the perfect gas (a theoretical construct) is exactly proportional to the thermodynamic temperature T (the one determined by TMD itself) and all gases at sufficiently low pressures behave as the perfect gas.

Before 1954 one defined the temperature scale using the gas thermometry (at $p \to 0$) as described above using the ice melting point and the water boiling point. Because in the laboratory practice it is easier to reproduce the water triple point, the definition has been changed and now the absolute temperature scale is fixed by ascribing to this point[10] (which corresponds to $p = 611.73$ Pa) the absolute temperature $T = 273.16$ K (exact value by definition). This has the effect (the value 273.16 has been chosen to get this!) that between the water boiling point and the ice melting point at 1 atm the temperature difference is (very nearly) 100 K and that these points correspond to 273.15 and 373.15 K, respectively (although only within some accuracy: more precise measurements may reveal small departures from these numerical values).

It should be noted that because the perfect gas scale relates directly to the thermodynamic temperature, the determination of temperature almost always is based on gas thermometry. In general, calibrating thermometers with respect to the absolute temperature scale is too long a story to be told here (it requires also developing the formal apparatus of thermodynamics).

Once the thermodynamic temperature is established, the commonly used (in Europe; in USA for example the Fahrenheit scale is in use) Celsius scale is defined as $t = T - 273.15$. On this scale the triple point corresponds to 0.01^oC. This approximately (to a quite good accuracy) coincides with the old Celsius temperature, which is now called the centigrade scale, defined by the Mercury in glass thermometer using the relation $h_{\rm Hg} = a_{\rm Hg}t_{\rm Hg} + b_{\rm Hg}$ and by ascribing 0^oC to the ice melting point (and therefore 100^oC to the water boiling point). The International Practical Temperature Scale is the set of accurately measured reference points plus a set of thermometers which should be used to interpolate between the reference points together with the interpolation procedures,

Gas thermometers are inconvenient and difficult to use when high accuracy is required, so they are used only to measure absolute temperature T. Other kinds of thermometers are used - the choice depends on the convenience and sensitivity required. To calibrate these other thermometers w.r.t. the absolute temperature a number of reference points have been measured very accurately. Among these are: the triple point of Hydrogen (13.81 K), triple point of Oxygen (54.361 K), the melting point of Zinc at 1 atm (692.73 K) melting point of Gold (1337.58 K).

Of the thermometers used in the laboratory practice the ones based on expansion of liquids (Mercury, Ethyl Alcohol, Pentane) cover the range of $-$hundred ^{o}C up to + a few hundreds ^{o}C.

Resistance thermometers are based on the variation of electrical resistance of a metal

---

[10] The precise definition of this reference point includes the condition that the isotopic composition of water must be that of the ocean water.

with temperature. They cover an even larger range of temperatures. E.g. thermometers using Platinum, which is easy to purify and has a rather high melting point (1770^{o}C) are highly accurate between 70 K and 1500 K.

Thermocouple thermometers use variation of the e.m.f. with the temperature. If one junction is kept at a fixed temperature, the e.m.f depends on the temperature of the other one. Using them requires measuring rather small voltages - this make this kind of thermometers difficult to work with if high accuracy is needed - but they can be miniaturized and respond quickly to changes of temperature.

Thermometers exploiting the conductivity of semiconductors. Current carriers must be thermally excited and the semiconductor conductivity is proportional to the temperature dependent factor $\exp(-\varepsilon/T)$. They are good thermometers from well below 1 K up to $\sim$ 600 K. At low T such thermometers have sensitivity of order 10^{-5} K and of order 10^{-3} K at room temperatures.

Thermometers based on carbon resistors are seful below 20 K. Below 10 K also have sensitivity of order 10^{-5} K. Below 5 K to somewhat below 1K liquid ^{4}He is used and between 1 K and 0.3 K - ^{3}He. In both cases T is found by measuring the vapour pressure. For yet lower temperatures one relies on paramagnetic salts. Their susceptibility goes like a/T (Curie law).

On the other extreme, above the Gold melting point only measurements of radiation emitted by hot bodies are used (radiation pyrometers).

Finally it should be stressed that the mere introduction of the notion of temperature - as **an indicator** of whether two bodies will be in thermal equilibrium, if they are brought to a contact through an diathermal wall - does not imply yet any correlation of its values with the sensation (experienced by our bodily senses for example) of warmth and coldness. Nothing as yet guarantees that higher (lower) t corresponds to what we feel as hotter (colder). Of course, one can arrange the perfect gas scale to reflect the degree of hotness but this cannot be demonstrated rigorously before *defining* the meaning of the terms “hotter” and “colder” operationally that is, in a way which is not based on our subjective physiological sensation. And this requires to investigate first what **heat** is.

LECTURE II (TMD)

Relying on the fundamental fact that (almost) every system, when isolated (adiathermally by rigid walls against influences from without, on which no other kind of work is performed by any means) sooner or later attains an equilibrium state in which no change is perceptible macroscopically and on 0TMDL (stated in two equivalent ways), we have inferred (by using a physical as well as a mathematical reasoning) the existence of a new (in addition to the system's all deformative parameters and to the single nondeformative one, like volume the V and pressure p in the case of simple fluids) state function characterizing every thermodynamical system - the empirical temperature t.

If the empirical temperatures t_A and t_B of two isolated systems A and B are equal, $t_A = t_B$, then no change in their states will be observed if they are brought into contact with one another through a rigid (preventing their mechanical contact and transfer of matter between them) diathermic wall. The two systems are then said to be in **thermal equilibrium** with one another. (Notice however, that the condition $t_A = t_B$ alone is not sufficient to ensure that these two systems are in **full thermodynamical equilibrium** with one another: this requires that in addition their pressures as well as chemical potentials - to be introduced in due course - be equal so that when the two systems are contacted through a movable wall which also permits transfer of matter between them, no change in their idividual states will observed). The empirical temperature plays therefore the role of an indicator of a possible thermal equilibrium between different systems. There is a huge arbitrariness in the definition of the empirical temperature t: a given t can be always replaced by $\tilde{t} = f(t)$ where $f(\cdot)$ is an arbitrary monotonic function. But once this arbitrariness in the definition of t has been fixed (by choosing a standard thermometric body), it is a single-valued function of the deformative parameters $X_1, \dots, X_{o-1}$ (like V), and the single nondeformative one y (like p):

$$t = t(X_1, \dots, X_{o-1}, y)\,, \tag{5}$$

(in general t depends also on the amount of matter in the system represented by the number of moles $n_1, \dots, n_r$ of its r chemical components - which are needed to fully characterize the equilibrium state when the system is not treated as a black box). Such a relation is called the **equation of state** of the system (of the body, of the substance). Naturally, thermodynamics by itself does not predict its form and it has to be determined experimentally (by measuring over some ranges of the relevant parameters various coefficients like $k_t = -(1/V)(\partial V/\partial p)_t$, $\alpha_p = (1/V)(\partial V/\partial t)_p$, $\beta_V = (1/p)(\partial p/\partial t)_V$, in the case of fluids and other substances) or else derived using the statistical physics approach.

We recall, however, that as yet no correlation between higher (lower) value of the empirical temperature and the subjective sensations of hotness (coldness) has been established. This requires defining precisely the notion of heat. We shall do it now, continuing to consequently develop thermodynamics as a phenomenological theory (as opposed to the Callenian thermodynamics which is - somewhat absurdly to my taste - constructed as a deductive theoretical system).

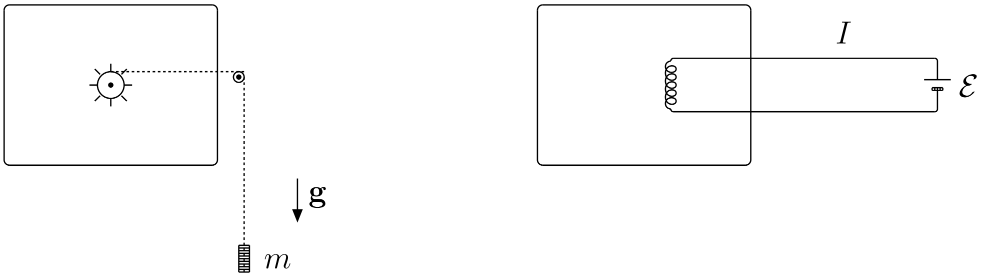

Figure 1: Two possible realizations of the historic Joule experiment.

Internal energy
We begin by introducing the notion of **internal energy**. Let us consider experiments similar to the historic one performed by Joule which originally was intended to measure the mechanical equivalent of heat. But for us the word "heat" has as yet no content, so we will interpret these experiments somewhat differently.

One realization of this sort of experiment (see the left panel of Fig. 1) can be the original Joule's paddle-wheel immersed in an adiathermally isolated (as ideally as it is possible) calorimeter containing a mass of a liquid (e.g. water or beer - recall who Jule was!). A measurable mechanical work can be performed on this system by rotating the wheel by an angle α applying to it a known couple D: $W = \alpha D$. Alternatively one can let a known mass m fall down the height h in the Earth's gravitational field $\mathbf{g}$ propelling the wheel: measuring its final speed $\mathbf{w}$ one gets the work $W = mgh - m\mathbf{w}^2/2$ (w neglect the kinetic energy of the wheel) done on the system (the liquid) by the device. It is observed that as a result of performing the work the state of the liquid has changed (its temperature has changed, its pressure has changed).

Alternatively, a resistive wire can be inserted in the calorimeter and a known current I passed through it during a period $\Delta\tau$ (as in the right panel of Fig. 1). If the potential difference across the wire is $\mathscr{E}$, the (electrical) work done on the liquid equals $W = \mathscr{E}I\Delta\tau$.

Similar experiments, employing different kinds of **directly measurable works** can be performed on the system (here the liquid). The important fact is that if the initial system's state is the same, always the same temperature change is obtained by the performance of the same amount of work. Pippard (whom I follow here) stresses that none of such experiments should be interpreted as transferring heat to the system: "So long as we take account only of what is observed, the deduction to be drawn from the experiment is (...):" *If a state of an otherwise isolated system is changed by the performance of work, the amount of work needed depends solely on the change accomplished, and not on the means by which the work is performed, nor on the intermediate stages through which the system passed between its initial and final states* (in the context of thermodynamics I would add here "equilibrium states" - see the footnote below).
This statement is 1TMDL as **applied to adiathermally isolated systems**.

This can be illustrated by a simple example. The state of a fluid (a gas or a liquid)

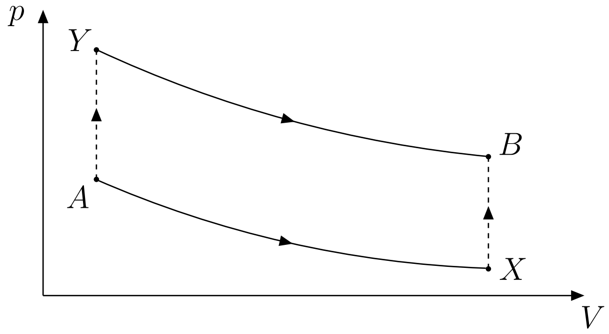

Figure 2: Two possible adiathermal processes taking a gas from its equilibrium state A to another equilibrium state B. The changes $A \to Y$ and $X \to B$ during which the system may not be in equilibrium (and strictly speaking its state cannot be then characterized by two parameters p and V) are marked by the dashed lines.

which is adiathermally isolated (against any uncontrolled influences from without) changes from A to B. Path 1 (Fig. 2): an amount of electric work is first performed isochorically (i.e. keeping the gas volume constant) on it, until it reaches the state Y; then the gas is adiathermally (quasistatically - see below) expanded performing some measured work and attains the state B; path 2: the gas is first adiathermally expanded reaching the state X and doing on the way some measured work and then an amount of work is done on it by, say, a paddle-wheel mechanism so that the gas attains the same final state B. Then the statement is that

$$W_{A\to Y\to B} = W_{A\to X\to B}\,.$$

The processes considered here need not be reversible (a word not defined yet) and at the intermediate stages the system may not be in equilibrium (parameters like pressure, temperature of the system may not be defined at these stages). It is only required that the system be adiathermally isolated, the works done on/by it measured, and the initial and final states be equilibrium states.[11]

It should be stressed that experiments of this kind (checking carefully that indeed the same work is needed to produce the same change) have never been really performed, most probably because of the rapid universal acceptance of the just stated form of 1TMDL. But "its manifold consequences are so well verified in practice that it should be considered to be established beyond any reasonable doubt" (Pippard again).

Relying on 1TMDL applied to adiathermally isolated systems one can now introduce the notion of **internal energy** U which by construction is a state function. If an adiathermally isolated system is brought from the state A to another state B by performing on it

---

[11] We require the states A and B to be equilibrium states because then they can be specified by giving the values of a few parameters only and only then the example is useful for illustration of thermodynamical considerations; the statement remains true also if one or both these states are not equilibrium states but it would then be harder if not impossible to be sure that the same state B has been reached on both paths.

an amount W of work, its internal energy is said to have changed by

$$\Delta U \equiv U_B - U_A = W \qquad \text{on adiathermal paths}\,. \tag{6}$$

1TMDL asserts that ΔU is determined by the states A and B only and not by the (adiathermal) path connecting these two states. So fixing for every system some reference state R and assigning to it (arbitrarily) an internal energy U_0, the internal energy of any other state A of this system is uniquely determined:

$$U_A = U_0 + W_{R\to A} \qquad \text{on adiathermal paths}\,. \tag{7}$$

In reality it may prove difficult to measure $W_{R\to A}$ directly but, again owing to 1TMDL, it can always be measured indirectly: a suitable roundabout path can in principle always be devised to get from a state R to another state A **or the other way around**, that is from A to R. In all textbooks it is remarked at this point that as a consequence of 2TMDL (which will be introduced later) a given path connecting the states A and B may not necessarily be traced in both directions (in given adiathermal conditions) but to determine ΔU it is sufficient that it can be traced in one way only. Since I have always had trouble to understand what this enigmatic statement is intended to mean, I advice the reader to think again on the paddle-wheel device: call the state R the state of the fluid when the mass m the fall of which propels the wheel is in the upper position and A the state of the fluid when the mass m is in the lower one. The path R to A can be realized (in fact in numerous ways). The path A to R evidently cannot and not only with the help of the paddle-wheel device but by any other means, so long as the liquid remains adiathermally isolated. It is clear that only transitions leading from a state of lower energy to a state of higher energy can be realized in this way.

In fact, it is not dificult to devise a system for which even this is not true. Take two identical amounts of the same fluid contained in two identical completely isolating vessels of volume V and form out of them a single compound system (which, hwever, possesses an internal adiathermal partition). If A is its state in which one fluid is at pressure p_1 and the other one at $p_2 > p_1$ and B is the state in which the first fluid has pressure p_2 and the other one p_1, it is clear that neither the transition $A \to B$ nor the transision $B \to A$ is possible (neither are all transitions to states in which the first system has $p_1' > p_1$ and the second one $p_2' < p_2$). This shows that ascribing energies to states of a system possessing an internal adiathermal partition is less direct - states of thermally homegeneous parts of such a system must be individually ascribed energies according to the procedure described above and energies of states of the formed compound system are then “more theoretical” constructs being **defined** as sums of energies of its separate parts.[12]

Although in general one cannot exclude that a similar situation cannot occur in the case of a system not possessing internal adiathermal partitions, one assumes (and this should be treated as an ancillary law) that if A and B are two states of a standard (i.e. with no internal partitions) system either a transition $A \to B$ or $B \to A$ is feasible

---

[12] This also shows that in formulating a phenomenological theory like thermodynamics one has to be”reasonable”...

adiathermally, at least as far as systems not possessing internal adiathermal partitions are concerned.[13]

In any case, what is important is that differences of internal energies in various states of a given system and, therefore, this quantity itself up to an additive constant can be determined by measuring **only mechanical or electrical works** and that the internal energy U is a function of state of that system, that is, it can be expressed (in the case of equilibrium states) as a function of the parameters needed to specify the state of the system: originally of the requisite deformative ones and of the single nondeformative one and then as a function of a chosen set of (secondary) parameters uniquely specifying the state. E.g.

$$\begin{aligned} U = U(V,p)\,, \quad &\text{or} \quad U = U(t,p) = U(V(p,t),\, p)\,, \\ &\text{or} \quad U = U(V,t) = U(V,\, p(V,t))\,, \end{aligned}$$

in the case of a simple fluid, and $U = U(X_1, \dots, X_{o-1}, y)$ or

$$U = U(X_1, \dots, X_{o-1}, t) = U(X_1, \dots, X_{o-1},\, y(t, X_1, \dots, X_{o-1}))\,,$$

etc. in the general case. In this way on the manifold of equilibrium states of every system (parametrized by some convenient set of state variables appropriate for that system) we can **superimpose the network of its internal energies**. Of course, microscopically this internal energy (by definition considered in the body's rest frame) consists of the kinetic and interaction energies of the system's microscopic constituents.

It also follows, since works can be added, that internal energy U is an additive quantity: if U_1 and U_2 are energies of two bodies taken separately, the internal energy U of these two bodies combined to form a compound system equals $U = U_1 + U_2$.

Heat as the "work defect" and 1TMDL (general)

Once a value of the internal energy U of a system is unambiguously (when the reference state R is chosen) assigned to every of its (equilibrium) states in the way sketched above, one can consider changes during which the system is **not** necessarily adiathermally isolated. It is then possible to accomplish a change of the system's state from A to B in more different ways and these new ways involve amounts of (measured) work which are different than $\Delta U = U_B - U_A$. For instance, in the considered example of a simple fluid taken from

---

[13] This can be formulated in a more economic way, which perhaps more clearly shows the essence of the problem and will help to see its relation to 2TMDL: if states of a system not possessing internal adiathermal partitions are characterized by $o-1$ deformative variables $X_1, \dots, X_{o-1}$, and a single nondeformative one y, then it is sufficient to accept as the ancillary law the statement that of any two states $P = P(X_1, \dots, X_{o-1}, y)$ and $P' = P(X_1, \dots, X_{o-1}, y')$ - these are the so-called isometric states, that is ones differing only by the value of the nondeformative variable y - either P' is reachable from P in adiathermal transitions, or the other way around (of course, in the course of the transition the variables $X_1, \dots, X_{o-1}$ may deviate from their initial and final values). The ancillary law adopted in the text then follows as a simple corollary. Moreover, if the internal energy U is taken for y, one can further assert that of the two, possible is this transition which leads to higher value of U. The connection with the example given in the text is then clear.

A to B the change from A to Y can be achieved by lightning a Bunsen burner under the (diathermal) flask containing the fluid and this certainly does not involve performing a work. So in this case $\Delta U = U_B - U_A \neq W_{A\to Y\to B}$. We then **define** the quantity Q by

$$Q = \Delta U - W_{A\to Y\to B}\,, \tag{8}$$

and call it (somewhat misleadingly, but this should cause no harm - we are already far from the discussions of the XIX-th century!) the **heat** taken by the system (the fluid in this case) on its way from A to B. More properly one should call Q **energy transferred to the fluid in the form of heat**. In other words - and this is my own great contribution to teaching thermodynamics! - **Heat is the "defect" of work** in obvious analogy to the binding energy of nuclei defined as the defect of mass in nuclear physics (nuclei weight less than protons and neutrons out of which they are composed and this deficit, called the "the mass defect", multiplied by c^2, is identified with their binding energy).

Heat defined in this way has all properties which are usually attributed to it (and which in the past supported treating heat as a kind of indestructible fluid flowing from one body to another): *i*) when absorbed or released by a body, it changes the state of this body (obvious - existence of diathermal walls proves that changes of states are not always due to work), *ii*) may be conveyed from one body to another by conduction, convection and radiation (obvious - even vacuum does not entirely inhibit changes effected by means other than mechanical or electric work - changes caused by radiation can be inhibited by walls perfectly reflecting electromagnetic waves), *iii*) in calorimetric experiments (in which one measures heat by the method of mixtures - this is what all problems in undergraduate physics courses are about: one puts into a calorimeter bodies which then exchange heat, or, saying it more properly, are in thermal contact with each other but are adiathermally isolated from the rest of the world, and one is asked - given the heat capacities of these bodies - to compute the final temperature or, given some other data, to compute the heat capacities etc.) heat is conserved if the bodies involved are placed in an adiathermally isolated vessel.

Only the property *iii*) requires a brief justification: consider a typical calorimetric experiment in which two bodies, 1 and 2, at different (empirical) temperatures $t_1 \neq t_2$ are brought into thermal contact in a calorimeter (in a Dewar vessel) the walls of which are adiathermal and rigid (no work can be performed on the bodies in the calorimeter by external agents). Therefore the total change ΔU of the system's internal energy must be zero

$$\Delta U = \Delta U_1 + \Delta U_2 = 0\,. \tag{9}$$

But

$$\Delta U_1 = W_1 + Q_1\,, \qquad \Delta U_2 = W_2 + Q_2\,, \tag{10}$$

and since the only work that could have been performed on the bodies was the work they performed on each other, so that $W_1 = -W_2$, one learns that

$$Q_1 + Q_2 = 0\,. \tag{11}$$

So, in calorimetric experiments (but not in general!) heat is conserved.

We have thus finally arrived at the **general 1TMDL** in the form[14]

$$\Delta U = W + Q\,. \tag{12}$$

In words: energy is conserved if heat Q is taken into account and recognized as a form of energy transferred from one system to another. Callen (whose celebrated textbook on thermodynamics will be exploited later on) describes the difference between energy transferred through a mechanical work and in the form of heat as results of couplings of external agents to different types of degrees of freedom of the system: mechanical work is due to coupling to external agents of the few globally defined macroscopic degrees of freedom (like the position of its center of mass, or volume) of the system, while a heat transfer occurs through a coupling of the microscopic degrees of freedom of the system and its surrounding.

The equivalence of heat and work can be illustrated by the example of a gas enclosed in a cylinder with a piston. The system is here the gas and the cylinder and the whole system is adiathermally isolated. If the gas expands doing some work $\bar{W}$ on the piston (not necessarily reversibly - we still do not know what this word means) and next the whole work $\bar{W}$ is converted into heat somehow (or just heat equivalent to this work is used) and added back to the system, its internal energy will return to the initial value (although its state will not in general be the initial one).

Notions "hotter", "colder" and their correlation with the scale of the temperature
We can now consider the problem of correlating the introduced empirical temperature with the properties of coldness and hotness. In calorimetric experiments like the one just considered, one body gains heat which the other body is loosing (considering experiments in which only two bodies take part). This is called **heat transfer** though this does not imply existence of heat as a physical substance (a "caloric" or a "flogiston") the movement of which from one body to another could be followed.

In general if any two bodies are brought into thermal contact in the conditions that no work is performed on either, a transfer of heat will occur (accompanied by the changes of states of both bodies) unless their temperatures are equal in which case they are in thermal equilibrium with one another. In the following reasoning important will be the fact, following from experience, that the **rate** of the heat transfer may usually be varied over a wide range, depending on the nature of the diathermal wall through which the bodies exchange heat (the rate is a measure of **thermal conductance** of the wall separating the bodies).

We adopt the definition (could one expect it to be different than this?!) that of the two bodies that one, call it A, which **loses heat** (negative Q_A) is **hotter** and the other body, call it B, which **gains heat** (positive Q_B) is by definition the **colder** one.

---

[14] Our convention is that Q and W will always stand for heat and work added **to** the system; heat and work extracted **from** the system will be denoted $\bar{Q}$ and $\bar{W}$, respectively. Of course, any one of these quantities can be either positive or negative.

Having introduced this definition we can now show that the hierarchy of hotness (coldness) defined in this way can be consistently correlated with the scale of temperature (the labeling of isotherms of bodies), that is, that the values of $t = t(V, p, \ldots)$, or $t = t(X_1, \ldots, X_{o-1}, y)$ in general, of the empirical temperature can be assigned to isotherms in such a way, that **all** bodies at temperature t_2 will be hotter (in the sense defined above) than **all** bodies at temperature t_1 **if** $t_2 > t_1$.

This is demonstrated by *reductio ad absurdum.* Let us assume this is not possible. Therefore it should be possible to find three bodies, A, B and C having temperatures t_A and $t_B = t_C$ (so B and C are in thermal equilibrium with one another) and yet such that A is hotter than B while C is hotter than A. For convenience we can assume that no work can be performed on these bodies and that they cannot perform work on each other either.

It is then possible to break somewhat the thermal equilibrium between B and C, varying slightly (almost infinitesimally) the temperature of B making it somewhat hotter than C but still colder that A (this should be possible because $t_A \neq t_B = t_C$; notice that we do not say whether t_A is higher or lower than $t_B = t_C$). If the three bodies are then brought into thermal contact, heat will flow (in agreement with the meaning of the words "hotter" and "colder") from A to B, from B to C and (because we have assumed C can from the beginning be hotter that A) from C to A. By adjusting the diathermal walls (their conduction rates) separating the bodies it would then be possible to establish a (dynamical) equilibrium of these three bodies. But this would **contradict** the converse of 0TMDL, for if any two of these bodies were separated from the third one, they would not be in thermal equilibrium.

The conclusion must, therefore, be that by virtue of 0TMDL **it is** possible to define the empirical temperature t so that if $t_2 > t_1$ than any body at t_2 is hotter than any body at t_1 and the established correlation between hotness (coldness) and the (empirical) temperature is generally valid for all (thermally homogeneous, i.e. having no internal adiathermal partitions) bodies and substances.

Since in this reasoning it was assumed that no work is performed on any of the bodies, it follows that the established correllation of hottness with the empirical temperature t implies that

$$\frac{\partial}{\partial t} U(X_1, \ldots, X_{o-1}, \, y(X_1, \ldots, X_{o-1}, t)) > 0 \, . \tag{13}$$

This means that the heat capacity at constant deformative parameters $X_1, \ldots, X_{o-1}$, defined as the ratio of the infinitesimal amount of heat needed to infinitesimally increase the body's (empirical) temperature t at fixed values of $X_1, \ldots, X_{o-1}$ to this temperature increase, $C_{X_1, \ldots, X_{o-1}} = \delta Q / \delta t = \delta U / \delta t$, is positive.

In the above reasoning used to demonstrate the possibility of establishing (by defining appropriately the scale of the empirical temperature t) the correlation between hotness and the temperature we have assumed that no work is performed on either of the bodies which exchange heat and this has led to (13). One can however relax somewhat these conditions admitting that the surrounding can do some work on each of them (but still assumiming they cannot do work on one another), provided those parameters which together with the temperature t are used to specify the states of each of these bodies remain constant when the bodies are brought into thermal contact (e.g. p instead of V when the rigid diathermal wall separating any of them from the surrounding is replaced by a movable one) - the correlation of t with the hotness will still obtain. The consequence of this is that absorption of heat by a body, the independent parameters of which other than its temperature t remain constant, will always cause an increase of the body's temperature t. Therefore the so called **principal heat capacities** (like C_V, C_p, C_M) characterizing a body are always **positive** (this will be important in establishing stability conditions of thermodynamic systems) by virtue of the operational definition of the notions "hotter" and "colder" and the (conventional) assignment of higher values of the (empirical) temperatures to the hotter bodies.

Reversible and irreversible changes (processes)

We have divided the changes which systems can undergo into adiathermal ones in the course of which the system is adiathermally isolated and ones in which the system is not isolated in this way. In this other, more general, kind of changes $\Delta U = W + Q$. We now inquire, under what conditions in an infinitesimal change of the system's state from one equilibrium state to another one the 1TMDL

$$dU = w + q\,, \tag{14}$$

in which dU is an infinitesimal change (which is an exact differential) of the system's internal energy and q and w are infinitesimal (elementary as one usually calls them in thermodynamics) heat and work,[15] can be written with w and q being forms (differential one-forms in thermodynamics called Pfaffian forms) on the space of the parameters $X_1, \dots, X_o$ (of which at least one is nondeformative - as we now have t and U as our disposal, they can be used to replace the original $o-1$ deformative and/or the single nondeformative variable; hence there may by now more than one nondeformative variable) characterizing **equilibrium states** of the considered system. Certainly, neither q nor w which represent infinitesimal heat and work added to the system in a change in which its internal energy changes by dU can be written as $df_Q(X_1, \dots, X_o)$ of $df_W(X_1, \dots, X_o)$, that is, as an exact differential (or a closed form) because in finite changes heat taken by the system and work done on it do depend on the path from the initial to the final state. We will first argue that under some well defined circumstances w can be written as an inexact form[16]

$$\hat{\omega}_W \equiv đW\,, \tag{15}$$

---

[15] We do not consider here the possibility that separately w and q are large but mutually cancel out leaving only an infinitesimal sum $w + q$.

[16] Thermodynamical tradition requires the differential forms which are not exact differentials of state functions be written with the slashed d, $đ$ (Kubo uses $d'W$), instead of $\hat{\omega}$ (which is the mathematical notation). In our Department prof. Cichocki was famous for assigning zero points to student works whenever he noticed the lack of the slash on d in the heat or work form...

which implies that in these conditions also q can be written as

$$\hat{\omega}_Q \equiv đQ\,, \tag{16}$$

simply because then $q = dU - đW$. Thus, if the necessary conditions are met, one will have the right to write

$$dU = đQ + đW\,. \tag{17}$$

But, as we are going to discuss in the next Lecture, owing to 2TMDL the character of the form $đQ$ is different than that of $đW$ and this has important physical consequences.

As usually, in order to understand the conditions which must be met it is convenient to consider a simple system, e.g. a gas enclosed in a cylinder fitted with a movable piston of cross section area A. In the equilibrium state the piston is at rest - the forces acting on it must be balanced so that the net force is zero. The forces at play are: a) the force which the gas exerts on the piston, equal $p\,A$, where p is the gas pressure, b) an external force $\mathbf{F}_{\rm ext}$ applied to the piston from outside (can be provided by the pressure of the surrounding or by other means), c) the static friction (it vanishes when the external force exactly counterbalances the gas pressure but in the presence of a friction this need not be so). When the piston starts to move, however slowly, dynamical frictional force enters into the play replacing the static one. Now suppose the piston has moved by $d\mathbf{x}$ while being acted upon by an external force $\mathbf{F}_{\rm ext}$. The work done by this force on the entire system (the gas, the cylinder and the piston) certainly is

$$\mathbf{F}_{\rm ext}\cdot d\mathbf{x} = -F_{\rm ext}dx = -\frac{F_{\rm ext}}{A}\,d(Ax) \equiv -p_{\rm ext}dV\,. \tag{18}$$

The work done on the system, assuming the absence of friction, would be this minus the kinetic energy acquired by the piston on the distance dx.

In general $p_{\rm ext}$ is not simply related to the pressure in the gas. Worse yet, if the piston moves quickly (or accelerates) the gas will not remain in equilibrium and it will not be possible to characterize its actual state by two parameters (p and V) only: the gas pressure will not be the same in the entire cylinder volume - it will be different from point to point. Only if the piston moves very slowly, without acquiring any acceleration (and only - if the walls of the cylinder are diathermal so that the gas is in thermal contact with the surrounding - if the heat transfer is realized reversibly - see the remarks below) can the gas inside the cylinder remain at any moment practically in equilibrium so that one can ascribe to it a unique pressure p (almost the same in the whole volume). Such a process will be called quasistatic. If in addition there is no friction, one can assume that $p_{\rm ext} = p$ and write the work done by the external force on the distance dx in the form

$$đW = -p\,dV\,, \qquad \text{or} \qquad đW = -p(V,t)\,dV\,, \qquad \text{or} \qquad đW = -p(V,U)\,dV\,, \tag{19}$$

depending on the choice of independent variables - (p,V), or (t,V) or (U,V) - characterizing the equilibrium states of the gas. In these conditions the work w becomes a well defined differential form on the space of the equilibrium parameters of the gas.

The equality $p_{\text{ext}} = p$ ensures also that the piston does not accelerate; if so, then it, strictly speaking cannot start to move - one is making here an idealization: the process in which one is allowed to write $đW = -p\,dV$ is the limiting case of a process which can be practically realized.

Thus, there are two conditions allowing to identify w with $đW(p, V)$:

- the change must occur vanishingly slowly - the process must be **quasistatic**, meaning that at each stage the system can be treated as if it were in full equilibrium[17] (within itself and with its surrounding)
- there should be no friction - the process must be **reversible** which in practical terms means that it occurs under an **infinitesimal** difference between p and p_{ext}; infinitesimal change of this difference (which reverses its sign) suffices to change the **direction** of the process

This can be further illustrated by the following considerations. If the piston is withdrawn suddenly, a rarefaction in the gas will occur and the work done by the gas will be smaller than if the pressure in the gas remained uniform. The extreme example is the process in which the piston is replaced by a wall impermeable to the gas particles which prevent the gas from diffusing into an additional volume of the cylinder. If this wall is removed the gas will fill the whole cylinder.[18] The removal of the wall can by done at no work cost at all. The gas will then not perform any work at all. And the pressure will not be defined for a while (the gas will be out of equilibrium). But if the additional volume is quite small, just an infinitesimal dV, the departure from equilibrium may be negligible and the pressure will remain well defined - the gas will practically stay in equilibrium - the expansion will be **quasistatic**. Now, the product $-p\,dV$ corresponding to such a change will certainly be non-zero. Yet it will not represent the work done on or by the gas - the work is exactly zero! One can also make in this way a finite change of the volume occupied by the gas (see Fig. 3), just by opening to it successively and always quasistatically, additional volumes dV until a finite change ΔV is accomplished. The work done by the gas will still be zero irrespectively of the fact that $\sum(-p\,dV) = -\int dV\,p \neq 0$. The process, while quasistatic, will not be reversible: one cannot change something infinitesimally to reverse its direction. This shows that quasistaticity does not imply reversibility although, as should be clear, any reversible process must be quasistatic.[19]

It is also instructive to consider the effects of the friction assuming that the piston is removed or inserted with vanishing velocity, without any acceleration and the equilibrium of the gas is maintained at each stage. When the gas is being compressed (the volume

---

[17] The practical realization of this postulate depends on the system; in the case of gases it is not very stringent - a uniform pressure establishes rather quickly in the entire volume occupied by the gas even if the piston moves at a nonnegligible speed.

[18] This process is called Joule process and will be discussed in more details in classes.

[19] As always, there is a lot of confusion in terminology here. Many authors put the equality sign between reversible and quasistatic processes, calling a process which is quasistatic in my sense but not reversible, a pseudostatic process. Somehow I cannot digest this semantic hair-splitting and maintain that linguistically it is much more natural to simply distinguish quasistaticity and reversibility.

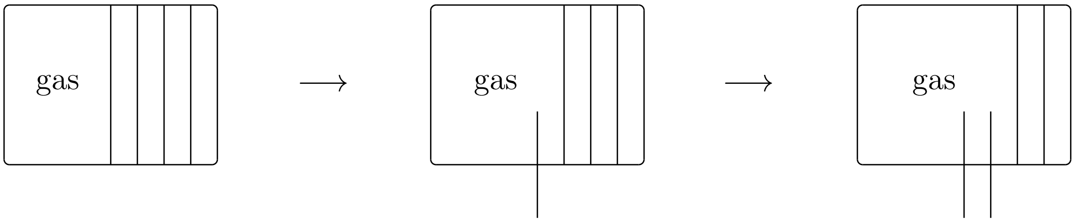

Figure 3: Successive openings to the gas additional infinitesimal volumes dV. The elementary works $w = 0$ despite that $p\,dV \neq 0$.

V decreases), $p_{\rm ext}$ must be greater than the pressure p of the gas, because the external force must overcome in addition the friction. When the piston is removed, $p_{\rm ext}$ must be smaller than p, because now it is the gas pressure which must overcome the friction. As a result of the complete cycle consisting of a compression followed by the decompression, the total work done by the external force will be positive (see Figure 4)

$$W = -\oint dV\, p_{\rm ext} > 0\,. \tag{20}$$

If the gas temperature at the end of this process is the same as in the initial state (because e.g. the whole cylinder is in thermal contact with a heat bath - a very large system having a fixed teperature; we assume that the volume after the cycle returns to its initial value), this means that the energy W given by (20) had to be somehow taken away from the system formed by the gas and the cylinder. If the final temperature is higher (assuming the cylinder does not absorb heat and does not conduct it) W is just the difference of the gas final and initial energies which agrees with our definition of internal energy U although in both situations we have to do with what conventionally - but incorrectly - is called conversion of work into heat (through the effect of the frictional force). But irrespectively of this, the whole process is in this case irreversible, for changing the direction of the piston requires a finite change (and not an infinitesimal one) of the applied external force. And certainly, even if the expansion-compression process was quasistatic, the work done on the system is not given by

$$-\oint dV\, p\,, \tag{21}$$

where p is the gas pressure, so w - the work done on the gas cannot be written as $-p\,dV$.

It should be also remarked at this place that **heat can be transferred reversibly** from one body to another one only if their temperatures differ infinitesimally, so that an infinitesimal change of the temperature of one of these bodies would reverse the direction of the heat flow. This is always the case when isothermic changes of a system are considered (provided other elements of such processes are carried out reversibly, there is no friction, etc.): one assumes that the heat which the system absorbs or rejects while undergoing such a change is exchanged with a reservoir (called also a heat bath) at the temperature

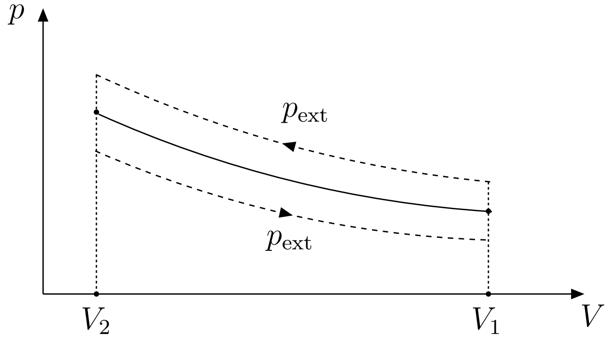

Figure 4: Compression and decompression of a gas in a cylinder fitted with a piston in presence of friction. Solid line represents internal pressure of the gas.

t_{res} equal to the system's temperature; owing to its very large size (infinite in the limit), the temperature of the reservoir stays unchanged (and its equilibrium in itself is not perturbed), now matter how big finite amount of heat it loses or absorbs.

Processes in which the system's temperature changes while they are exchanging heat (e.g. isobaric, that is occurring at constant pressure, expansion of a gas) can be treated as reversible provided one imagines them as split into small subprocesses in the course of which the system's temperature can be treated as approximately constant and the system is successively contacted with a sequence of reservoirs having temperatures adjusted to the actual system's temperature.

On the other hand, while every heat transfer occurring between two bodies at temperatures differing by a finite amount can always be imagined to occur **quasistatically**, so that the bodies involved can be treated as remaining practically in equilibrium at every stage (it is sufficient to imagine that the heat transfer is realized through a diatermal wall of nonvanishing, but arbitrarily low thermal conductivity, or it is realized in small portions just by temporarily breaking the thermal contact between bodies - this is similar to the already discussed gas expansion into the vacuum which, despite being irreversible, can be realized quasistatically) - such a process is always irreversible

Thus when the fluid undergoes a reversible change, one can write

$$dU = q - p\,dV\,. \tag{22}$$

This means that in such a change also q can be written as $đQ$, that is, as a differential form on the space of the systems' parameters, although at the moment we do not know how to do it in the way other than $dU + p\,dV$. Changing this situation requires 2TMDL.

Other types of reversible works

As said in the introductory part of these lectures, thermodynamics can be applied to very different physical systems on which different kinds of works can be done. Here we list some of them to prepare the ground for problems which will be assigned as class, home and colloquia works.

- Work required to enlarge the soap film spanned on a frame (picture).

$$đW = \gamma\,d\mathcal{A}\,, \tag{23}$$

where $\mathcal{A}$ is the area of the film and γ its surface tension. The relation $\gamma = \gamma(t, \mathcal{A})$ plays here the role of the equation of state. Usually γ depends on temperature only. Again, the change of the area must be made slowly to maintain equilibrium; friction is almost absent here, so a quasistatic change of $\mathcal{A}$ is almost automatically reversible.

- Work needed to deform a rubber band (many nice problems with this systems can be formulated) is

  $$\text{đ}W = K dL\,. \tag{24}$$

  L is here the rubber length and K its strain. The relation $K = K(t, L)$ plays the role of the equation of state. Reversibility requires that the deformation of the rubber band be fully elastic that is the rubber should not exhibit hysteresis.

- Work needed to stretch a thin wire.

  $$\text{đ}W = K dL\,, \tag{25}$$

  where L is the wire length and L its tension. This can be reversible if the deformation is fully elastic (and not plastic; again no hysteresis is allowed). The relation $K = K(t, L)$ plays the role of the equation of state of the wire. If the deformation is small it takes the form of the Hooke's law $K = k(T)(L - L_0)$, where L_0 is the length of the unstretched wire.

- The preceding example can be generalized to elastic deformations of a solid. In this case one considers the local displacement vector $\mathbf{u}(\mathbf{x}) = \mathbf{x}'(\mathbf{x}) - \mathbf{x}$. The changes of of distances in the body after the deformation are encoded in the tensor

  $$\mathbf{u}^{ij} = \frac{1}{2}\left(\frac{\partial \mathbf{u}^i}{\partial \mathbf{x}^j} + \frac{\partial \mathbf{u}^j}{\partial \mathbf{x}^i} + \sum_l \frac{\partial \mathbf{u}^l}{\partial \mathbf{x}^i}\frac{\partial \mathbf{u}^l}{\partial \mathbf{x}^j}\right) \approx \frac{1}{2}\left(\frac{\partial \mathbf{u}^i}{\partial \mathbf{x}^j} + \frac{\partial \mathbf{u}^j}{\partial \mathbf{x}^i}\right), \tag{26}$$

  in terms of which $d\ell'^2 = d\ell^2 + 2\mathbf{u}^{ij} d\mathbf{x}^i d\mathbf{x}^j$. The eigenvalues $\lambda^{(1)}(\mathbf{x})$, $\lambda^{(2)}(\mathbf{x})$, $\lambda^{(3)}(\mathbf{x})$ of the the tensor $\mathbf{u}^{ij}(\mathbf{x})$ determine local changes of the volume:

  $$dV' = dV\left(1 + \mathrm{tr}(\mathbf{u}^{ij})\right) = dV\left(1 + \lambda^{(1)}(\mathbf{x}) + \lambda^{(2)}(\mathbf{x}) + \lambda^{(3)}(\mathbf{x})\right). \tag{27}$$

  The elementary work done on the solid when it is elastically (not plastically!) deformed reversibly is given by

  $$\text{đ}W = \int_V d^3\mathbf{x}\, \boldsymbol{\sigma}^{ij} d\mathbf{u}^{ij}\,, \tag{28}$$

  where $\boldsymbol{\sigma}^{ij}$ is the stress tensor. The Hooke's law (the equation of state) takes in this case the form

  $$\mathbf{u}^{ij} = \frac{1}{9K}\boldsymbol{\delta}^{ij}\boldsymbol{\sigma}^{ll} + \frac{1}{2\mu}\left(\boldsymbol{\sigma}^{ij} - \frac{1}{3}\boldsymbol{\sigma}^{ij}\boldsymbol{\sigma}^{ll}\right) \tag{29}$$

or

$$\boldsymbol{\sigma}^{ij} = K\mathbf{u}^{ll}\,\boldsymbol{\delta}^{ij} + 2\mu\left(\mathbf{u}^{ij} - \frac{1}{3}\,\mathbf{u}^{ll}\,\boldsymbol{\delta}^{ij}\right). \tag{30}$$

where K is the compressibility modulus and μ is the shear modulus. Knowing their temperature dependence is equivalent to knowing the equation of state. In the case of a fluid under a hydrostatic pressure p the stress tensor takes the form $\boldsymbol{\sigma}^{ij} = -p\boldsymbol{\delta}^{ij}$ and the work reduces to $-p\,dV$. (See L&L, vol. 7, Elastomechanics.)

- Work needed for magnetization. Magnetic materials (paramagnetic or diamagnetic) put in an external magnetic field of strength $\boldsymbol{\mathcal{H}}_0$ (the subscript 0 indicates this is the magnetic field produced by the currents, in a coil, for example) acquire a local magnetization $\boldsymbol{\mathcal{M}}(\mathbf{x})$ and the work (which a battery must provide) needed to magnetize a given magnetic body is (using the illegal Gauss system of units and assuming a magnetic field strength homogeneous in space)

  $$đW = \boldsymbol{\mathcal{H}}_0\!\cdot\! d\mathbf{M}\,, \tag{31}$$

  where $\mathbf{M} = \int d^3\mathbf{x}\,\boldsymbol{\mathcal{M}}$ is the total magnetization of the body. The relation $\mathbf{M} = \alpha(t, p, \boldsymbol{\mathcal{H}}_0)\boldsymbol{\mathcal{H}}_0$ plays here the role of the equation of state. In special cases one can replace $\mathbf{M}$ by $V\boldsymbol{\mathcal{M}}$ and write equation of state as $\boldsymbol{\mathcal{M}} = \chi(t)\boldsymbol{\mathcal{H}}_0$ with $\chi(t)$ being the magnetic susceptibility of the material. Typically (if the temperature is not too low) $\chi(t) = a/t$, $a =$const. (Curie law).

- Work needed for polarization. A dielectric material placed in a uniform external electric field of strength $\boldsymbol{\mathcal{E}}$ acquires a local polarization $\boldsymbol{\mathcal{P}}$ and the work needed to polarize a piece of a dielectric material is

  $$đW = \boldsymbol{\mathcal{E}}\!\cdot\! d\mathbf{P}\,, \tag{32}$$

  where $\mathbf{P} = \int d^3\mathbf{x}\,\boldsymbol{\mathcal{P}}$. Again the (temperature-dependent) relation between $\mathbf{P}$ and $\boldsymbol{\mathcal{E}}$ plays the role of the equation of state.

In general we will write the infinitesimal works performed reversibly on a thermodynamic system (simple or non-simple) as

$$đW = \sum_{i=1}^{o-1} y_i(X_1, \dots, X_{o-1}, y)\, dX_i\,, \tag{33}$$

denoting by y_i the generalized “forces” ($y_1 = -y = -p$ in the case of a simple fluid).

A word concerning the mentioned hysteresis is perhaps in order here. It happens when the system under the same external conditions traces different paths when these external conditions are (slowly) changed. A good example is provided by isothermally magnetizing a specimen of iron by switching on an external magnetic field $\boldsymbol{\mathcal{H}}_0 = \mathcal{H}_0\mathbf{e}$. If the field is turned on slowly, the specimen passes through states which can be marked as equilibrium states on the indicator diagram (on the (M, $\mathcal{H}_0$) plane). Yet when the field is switched off the specimen of iron does not return to its initial state - a nonzero magnetization remains. In fact, at fixed temperature t and a given value of $\mathcal{H}_0$ there is a unique magnetization of the specimen corresponding to its full equilibrium but when the field is switched on for the first time, the process of reaching the equilibrium is very slow and the specimen is rather in a metastable state from which it would reach the true equilibrium only after a very long time. So the process of magnetizing a ferromagnetic material is not really quasistatic in the sense that the system does not necessarily passes through a sequence of true equilibrium states.

A digression
Before we move further on, it is useful to make a short summary. We have started by characterizing equilibrium states of any thermodynamic system by the deformative parameters of the (generalized) geometric nature - directly related to works which can be performed reversibly on the system and a single nondeformative parameter. In the case of a fluid these are V and/or $\mathbf{M}$ and/or $\mathbf{P}$ if systems with nontrivial electromagnetic properties are considered and pressure p. These parameters are said to be state functions. This means that (equilibrium) states of the system can be viewed (and this is, according to my experience, the most convenient view to adopt in thermodynamics, although one can develop more sophisticated mathematically points of view) as forming an abstract manifold (in the mathematical sense of this term) on which the functions like volume, pressure, etc. are defined. Next we have introduced some additional quantities: the empirical temperature t and the internal energy U (there will be also others). These too should be treated as functions defined on the manifold of system's states. A manifold, to be explored, requires introducing on it a system of coordinates which, mathematically speaking, map its points onto a space $\mathbb{R}^o$ (or $\mathbb{R}^{o+r}$, if the system is not treated as a black box and the numbers $n_1, \dots, n_r$ of moles of its material constituents are relevant). A coordinate system is therefore a set of o independent functions (their number o equals to, or rather defines, the dimension of the manifold) defined on the manifold. Thus p, V, or p, t, etc. should be treated as coordinates on the manifold of the system states. Once the system of coordinates has been chosen, other functions like U, t or S (entropy - to be introduced later) become functions of the coordinates used to identify points on the manifold. It should, however, be clear that the division into coordinated and functions is not fixed once for ever: in some applications what formerly was treated as a function can now be treated as one of the coordinates. It is known from the theory of manifolds that some systems of coordinates may be ill defined in some regions of the manifold[20] - the same can happen - si puo dare (compressed to "podarsi" in modern italian) as nostri fratelli Italiani would say - in some cases in thermodynamics: apparent paradoxes, contradicting e.g. 2TMDL, are mostly problems related to the wrong choice of coordinates on the manifold of states and not real paradoxes. Furthermore, differential forms $đW$, $đQ$

---

[20] Recall in this connection that the familiar spherical coordinates (θ, φ) on the sphere are ill defined at the poles (when $\theta = 0$ or π, the value of φ is arbitrary).

discussed here, which in fact are, like vectors or tensors, geometric objects, should also be viewed as forms defined over the manifold of the system's states and as such can always be expressed in the chosen coordinate system.

Other state functions
Already here one can introduce other state functions although the roles they play in thermodynamics cannot be appreciated at this stage (here these functions are introduced purely formally). In the case of a simple fluid one frequently uses **enthalpy** $H = U + pV$, **Helmholtz free energy** $F = U - TS$ or **Gibbs function** $G = U - TS + p\,V$ called also **free enthalpy**, where both:[21] the entropy S and the absolute thermodynamical temperature T - a function of the empirical temperature - will be introduced in Lecture IV as the consequences of 2TMDL. The introduced functions are called **thermodynamic potentials** (also U and entropy S are called thermodynamic potentials) for in specific situations they indeed play roles analogous to the role the potentials $V(q^1, \dots, q^s)$ play in Mechanics - they determine equilibrium states of thermodynamical systems. Of course, more such potentials (their names are not rigorously codified) can be formed in the case of nonsimple systems characterized by more than two parameters. As state functions, all of them can be expressed in terms of any set of independent state parameters.[22]

Let us briefly demonstrate the usefulness of enthalpy in chemistry. Most of chemical reactions occur at constant (atmospheric) pressure. During such a reaction the system, consisting of reacting substances (different at the end and initially) which change their volume in the course of the reaction, does work on the surrounding atmosphere (against its pressure $p_{\rm ext}$. If the initial and final states of the system are in equilibrium within themselves and with the surrounding atmosphere ($p = p_{\rm ext}$, $t = t_{\rm ext}$), 1TMDL can be written as $U_{\rm fin} = U_{\rm in} + W + Q$, with W the work done by the atmosphere on the system and Q the heat absorbed by the system. Since the work done by the surrounding on the system is $-p_{\rm ext}(V_{\rm fin} - V_{\rm in})$, and in the initial and final states $p = p_{\rm ext}$, one can write

$$U_{\rm fin} = U_{\rm in} - p\,(V_{\rm fin} - V_{\rm in}) + Q\,, \tag{34}$$

or

$$(U_{\rm in} + p\,V_{\rm in}) - (U_{\rm fin} + p\,V_{\rm fin}) \equiv H_{\rm in} - H_{\rm fin} = \bar{Q}\,. \tag{35}$$

Thus the heat $\bar{Q}$ released in a chemical reaction occurring at constant pressure is given by the difference of enthalpies of the initial and final substances. Notice that it is not required that the thermodynamic equilibrium be maintained during the reaction; only the initial and final states must be equilibrium states. (Examples in classes.)

Enthalpy, as will be discussed in classes, is conserved in the so-called Joule-Kelvin process in which a gas passes irreversibly from a state in which its temperature and

---

[21] Since both: U and the product pV have the physical dimension of energy, also the product TS must have the same dimension; the individual dimensions of T and S can only be fixed by adopting some convention.

[22] Later we will see, however, that each of these thermodynamic potentials has a preferred, or "natural" (in the sense that will be elucidated) set of its variables.

pressure equal t_1 and p_1, respectively, to another state in which its pressure p_2 is lower and temperature equals t_2.

Finally, enthalpy is conserved also in various flow processes (that is when one has to do with a steady state rather than with an equilibrium state of the considered system) - its conservation generalizes to compressible fluids the Bernoulli law (which applies only to incompressible fluids).

LECTURE III (TMD)

We have already introduced and discussed two of the four laws of thermodynamics. In the preceding Lecture it was argued that in infinitesimal **reversible changes** it is possible to represent the elementary work w in the generally valid form $dU = q + w$ of 1TMDL as a differential form $đW$ (or $đw$, if changes du of the molar internal energies are considered) defined on the space of parameters characterizing equilibrium states of the considered system. In such changes, since dU and $đW$ are well defined differential forms, also q must be a differential form $đQ$. However, on the basis of what has been done, there is no way to write this differential form differently than $đQ = dU - đW$, for example, considering a simple fluid, we can only write

$$đQ = dU(p, V) + p\,dV, \tag{36}$$

if the variables V and p are taken for independent ones (V as the deformative one). The heat Q taken by the system in a finite process (a finite change) in which it reversibly passes from the equilibrium state A to another equilibrium state B is then given by

$$Q = \int_A^B đQ = \int_A^B (dU(p, V) + p\,dV)\,, \tag{37}$$

the integral being taken along the curve representing the process on the (V, p) plane.

Adiabatic changes of the system, that is, its **reversible adiathermal** changes (in my terminology), are changes in which the form $đQ$ vanishes. More mathematically, $đQ$ projected onto paths representing such processes in the space of parameters is zero - it gives zero on all vectors tangent to such paths. Adiabats of a simple fluid are therefore curves in the (V, p) or (V, t) or (t, p) spaces determined by the solutions of the differential equation

$$dU + p\,dV = 0\,, \tag{38}$$

written in the set of variables, (V, p) or (V, t) or (t, p), which has been chosen (on the basis of convenience) to work with, and an initial point (initial state). In the general case of a system (treated as a black box) the equilibrium states of which are determined by the parameters $X_1, \dots, X_{o-1}, y$ and for which, as in (33), y_i are the generalized forces with which the system can oppose itself to the external actions on it, adiabats are all curves (paths) onto which projections of the one-form[23]

$$đQ = dU(X_1, \dots, X_{o-1}, y) - \sum_{i=1}^{o-1} y_i(X_1, \dots, X_{o-1}, y)\,dX_i\,, \tag{39}$$

---

[23] Recall, that the number o of independent parameters equals 1 plus the number of works which can be reversibly done on the system, or 1 plus the number of the deformative variables needed to characterize the system.

vanish. Of course, if $o > 2$, i.e. if the system is not simple, the initial point does not determine the path uniquely: the condition $đQ = 0$ provides only one constraint on o increments ($o - 1$ increments dX_i and one dy) and therefore from the initial point (and all subsequent ones) one can "adiabatically walk" in many different directions.

Unfortunately in practice one usually does not know a priori U as a function of the system's parameters like V and p, or $X_1, \ldots, X_{o-1}, y$ in the general case - the assignment of the internal energies to different states of the system is an example of the typical "paper and pencil" theoretical construction! - and therefore one cannot go too far in this way with solving various thermodynamical problems.[24]

At this point it is amusing to tell the reader that the system called "perfect or ideal gas" (and the "perfect magnetic material") has been invented partly in order to have a system on which to torment students. Indeed, including as part of its definition the information that the perfect gas internal energy U (treated as a function of V and t) is independent of its volume[25] V, the main obstacle for playing with this system is removed, because passing to V and t as the independent variables, the heat form (36) of the perfect gas can be explicitly written as

$$đQ = \left(\frac{\partial U}{\partial t}\right)_V dt + \left[\left(\frac{\partial U}{\partial V}\right)_t + p(t,V)\right] dV = C_V^{(t)}\, dt + p(t,V)\, dV. \tag{40}$$

$C_V^{(t)}$ is here the heat capacity of the system i.e. the heat absorbed by it when its empirical temperature t changes by one unit at constant volume V, but since usually in the definition one includes also the constancy of $C_V^{(t)}$ and, moreover, since the empirical temperature defined by the perfect gas equation of state happens to be (proportional to) the absolute temperature T (to be introduced in this Lecture), the fact that one (consciously or unconsciously) replaces t by T has no consequences.

The adopted definition of the perfect gas allows to solve problems of the sort "what heat the gas absorbs when it isothermally and reversibly expands from the pressure p_1 to $p_2 < p_1$": since the internal energy of the perfect gas depends on the temperature only, it stays constant during the isothermal expansion and, by 1TMDL, $Q = -W = \bar{W}$, where

$$\bar{W} = -\int_\Gamma đW = \int_\Gamma p\, dV(T,p) = -nRT \int_{p_1}^{p_2} \frac{dp}{p} = nRT \ln(p_1/p_2)\,,$$

upon using the equation of state in the form $V(T,p) = nRT/p$. The same integral gives also the answer directly because as can be seen from (40), on isotherms, owing

---

[24] However, some of the thermodynamical relations **can** be obtained with this limited knowledge which is encoded in 0&1TMDLs - see the home and colloquia problems to this course.

[25] This is usually backed by the physical argument that molecules of the perfect gas are mutually noninteracting and therefore the internal energy of the gas is just the sum of kinetic energies of individual molecules; the volume V occupied by the gas determines only the relative positions of the molecules but since there is no contribution to U of the interaction energies (which would depend on the relative distances of the gas molecules), changes of the volume do not have in this idealized case any impact on U. Of course such argumentation goes beyond the framework of the strict classical thermodynamics.

to the assumption that $(\partial U/\partial V)_t = 0$, the heat form of the perfect gas takes the form $đQ = p\,dV$. Such a simple reasoning would not be true if the gas satisfied e.g. the Van der Waals (VdW in short) equation of state, because, as it will be possible to show by appealing to 2TMDL, the internal energy U of this system does depend on its volume V.

As a matter of facts, as a consequence of 2TMDL (i.e. of the special character of the form $đQ$), the dependence of the system's internal energy U on the volume V is directly determined by this system's equation of state and it could happen that the assumption of independence of the internal energy U on the volume V is (mathematically) inconsistent with the perfect gas equation of state $p\,V \propto t \propto T$ (fortunately, it is consistent, as you will be able to check).

Furthermore, in reality, the usually assumed independence of the heat capacity C_V of the temperature (2TMDL does not constrain the dependence of C_V on the temperature) is (approximately) true only if the perfect gas - let us stick to this idealization - is a monoatomic one and structurless[26] - $C_V = nc_v$ in this case equals $\frac{3}{2}nR$, where n is the number of moles. Gases, the molecules of which are composed of more than one atom (the majority of real gases) have heat capacities only (to a good degree of accuracy) piecewise constant: at almost all temperatures at which gases exist as gases (and not as liquids into which they eventually change when the temperature is lowered) the three-dimensional motion of gas molecules as a whole (of their centers of mass) is quasiclassical (can be represented as in classical mechanics) and contributes $\frac{3}{2}R$ to C_V/n; however owing to the principles of quantum mechanics which must be employed to properly treat the internal motions (rotations and vibrations) of compound molecules, the molar heat capacities C_V/n of multiatomic gases rise starting from some temperatures to $3R$ ($\frac{5}{2}R$ if the gas molecules are composed of two atoms only) and then to even higher values depending on the number of the vibrational degrees of freedom of the molecule ($+1R$ per each vibrational degree of freedom; the number of vibrational degrees of freedom of a molecule equals $-3-3+3\times$the number of atoms, or $-3-2+3\times 2 = 1$ in case of two atoms). All this, as we will see, can be predicted within the statistical physics approach (in phenomenological thermodynamics molar heat capacities together with their dependence on temperature must be taken directly from measurements or else can be related to other quantities taken from experiment) - the gas of mutually noninteracting molecules is one of a few completely solvable problems (see Lecture XII). Within the statistical physics approach it will also become clear that the temperatures T at which the rise of the molar heat capacity occurs can be estimated from the simple rule

$$T \sim E_{\rm exc}/k_{\rm B}\,, \tag{41}$$

where $E_{\rm exc}$ are typical energies of the rotational and vibrational excitations of the gas molecule and $k_{\rm B}$ is the Boltzmann constant,[27] $k_{\rm B} = 8.617\times 10^{-5}$ eV/K. Typical energies of rotational excitations of molecules lie in the $3\times 10^{-5} \div 5\times 10^{-3}$ eV range while those of

---

[26] Even the heat capacity of monoatomic gases can exhibit local peaks if their electronic lowest energy states are split due to the spin-orbit or hyperfine interactions.

[27] These are the right units in which the Boltzmann constant should be remembered; giving it in J/K is as useless as measuring the Warsaw-New York distance in microns or atomic sizes in parsecs...

vibrational excitations in the range $0.1 \div 1$ eV (energies of the rotational excitations are lower than energies of the vibrational ones - a simple physical intuition says that it is easier to rotate something than to make it vibrate), so typical temperatures at which C_V rises, are in the region of a few to hundreds Kelvins (excitations of rotations) and then in the range of tens to hundreds thousands Kelvins; in the intermediate temperatures there can also be some transient departures (also predictable within the statistical approach; these are the "anomalies" mentioned in the introductory part of Lecture I) from the constancy of specific heats due to e.g. a fine structure of the atomic ground state (the corresponding energies are of order 10^{-2} eV). Thus, at least some quantum effects manifest themselves in the specific heats of gases only at rather high (compared to the room ones) temperatures (not only, as one naively could think, at very low temperatures)!

2TMDL
Only some of all changes of thermodynamical systems permitted by 1TMDL, that is, by the conservation of energy, are observed to occur in the real world. All changes have a clear tendency to occur preferentially in one direction (not in both); to take the simplest phenomena: if the mechanical energy is lost as a result of friction or viscosity it cannot be recovered (without other changes), chemical reactions occur evidently irreversibly, mixing of different gases is also irreversible, etc. The preferred direction of the change is perhaps most clearly manifested in the distinction between a hot and a cold body: although 1TMDL does not forbid the opposite (so long as the total energy is conserved), heat flows between two bodies which initially are not in thermal equilibrium in such a direction as to eventually bring them into equilibrium (equalize their temperatures). This allowed us in Lecture II to define operationally which of the two bodies is hotter and which one is colder. In fact, the basic assumption underlying the whole phenomenological thermodynamics (but also the equilibrium statistical physics) is that systems left to themselves eventually attain an equilibrium state; reversion to the original state is never observed (if it happened, the notion of equilibrium would loose any sense).[28]

Although in the preceding Lecture we found it useful to single out reversible changes which can occur in both directions - only in such processes the form $đW$ can be used as representing the work done on the system - they are only theoretically useful idealizations - they require stringent conditions, impossible to fulfill in practice (like $p = p_{\rm ext}$, $t = t_{\rm ext}$, quasistaticity etc.). Normal and prevailing type of behaviour of real systems are their

---

[28] It is perhaps fair to say already in this place that looking at the system from the microscopic point of view it may seem that returns of (macroscopic) systems to states macroscopically indistinguishable from the initial ones are not forbidden by the fundamental laws (of classical or quantum mechanics) but are only very improbable. So it may seem that from the microscopic perspective the discussed tendency of changes to occur in only one direction is only a statistical effect. Of course, probabilities of such returns are so fantastically tiny (one talks here of numbers as small as $10^{-10^{n-\rm teen}}$) that they never happen in practice and are never observed and, therefore, phenomenological thermodynamics which is based on what is really (and not what hypothetically could be) observed can safely rest on the discussed assumption. More importantly, however, this is rather the question of the definition within the statistical approach of the macroscopic equilibrium state of a thermodynamic system - as will be discussed, the proper one allows for such fluctuations and takes them into account.

irreversible changes.

These obvious (so obvious from the everyday experience that it took longer to accept 1TMDL - accepting that the mechanical energy is not lost but gets converted into the internal energy required complicated quantitative measurements - than to accept 2TMDL) observations underlie 2TMDL which generalizes them and promotes to the universal law of Physics. There are various formulations of this law and in view of the fundamental role it plays in physics, we will discuss all of them (although not all in equal depth).

2TMDL (R. Clausius, 1850, in Pippard's words):
It is impossible to devise an engine which, working in a cycle, shall produce no effect other than the transfer of heat from a colder body to a hotter body.

Of course, we have *defined* the notions of "colder" and "hotter" on the basis of the direction the heat flows, so this statement of 2TMDL may appear somewhat tautological, but defining these notions we had in mind only a direct contact of two bodies through a diathermal wall; the Clausius' principle says that reversing this direction is impossible at all, by any means and by using any roundabout physical process.

The crucial clause in this (and in the Kelvin's one given below) formulation is "working in a cycle": It is possible, for example, to expand a gas isothermally in contact with a cold body, so that it absorbs some heat Q, then to compress it adiathermally making it hotter, and then to bring it into contact with a hotter body (reservoir) and, compressing it further isothermally transfer heat to the hotter body (making the two necessary compressions at the cost of the work gained at the first stage). Such a process does not violate 2TMDL as formulated above for it is not a cycle - at the end the gas will be in a different state than initially. Only if it were possible to bring the gas back to its original state without undoing the heat transfer (and without introducing changes in the surrounding), could violation of the Clausius' 2TMDL be claimed. Only cyclicity can guarantee that the process would be repeatable and could serve to transfer an arbitrary amount of heat from a colder body to a hotter one.

2TMDL (Kelvin, 1851, again in Pippard's words):
It is impossible to devise an engine which, working in a cycle, would produce no effect other than the extraction of heat from a reservoir and performance of an equivalent amount of mechanical work.[29]

In contrast to the Clausius' formulation, which as all German philosophy of that time (Hegel, Kant & Schelling!) concentrates on rather abstract notions, the one of Kelvin clearly bears an imprint of an utterly pragmatic approach, characteristic of the times of the British industrial revolution: in simple words it communicated to engineers and inventors of machines what they should not hope to achieve. Its alternative formulation can read:

---

[29] It is perhaps worth stressing that the words "reservoir" here and "body" in the Clausius' formulation of 2TMDL should be understood as meaning a thermodynamic system which, even if compound, is thermally homogeneous.

It is impossible to absorb heat from a reservoir and to convert it all into mechanical work without introducing any other changes in the system and its surrounding.

Again, by expanding isothermally a gas remaining in thermal contact with a reservoir of heat (the reservoir's internal energy) it is possible to perform some work using up the extracted heat but such a process is not cyclical and therefore, it does not contradict the Kelvin's 2TMDL.

The formulations of 2TMDL by Clausius and Kelvin are easily proven to be equivalent. This is a sort of a scholastic exercise, which I recall quickly here.

1. Clausius⟶Kelvin. We argue that if the Kelvin's principle were violated, one could violate also the Clausius one (proof by *reductio ad absurdum* - *a.a.*). Suppose there are two reservoirs (bodies), a colder one at t_1 and a hotter one at $t_2 > t_1$. Suppose Kelvin is wrong, and a heat $\bar{Q}$ can be taken from the colder reservoir and all converted into work $\bar{W}$. Then this work can be used up to run a reversible Carnot engine[30] which takes heat $\bar{Q}_1$ from the reservoir at t_1 and transfers the positive amount $Q_2 = \bar{Q}_1 + \bar{W}$ of heat to the reservoir at t_2.

2. Kelvin⟶Clausius. We argue that if the Clausius' principle were violated, one could violate also the Kelvin's one (again *a.a.*). If Clausius were wrong, one could transfer a heat $\bar{Q}_1$ from the reservoir at t_1 to another one at $t_2 > t_1$ and then run a Carnot engine between the two reservoirs; it could be arranged so as to take the heat $\bar{Q}_2 > \bar{Q}_1$ from the reservoir at t_2, give back the heat $\bar{Q}_1$ to the reservoir at t_1 producing the net work $\bar{W} = \bar{Q}_2 - \bar{Q}_1$, in effect entirely at the cost of heat taken entirely from the reservoir at t_2.

There is one more formulation of 2TMDL, due to Carathéodory, which, compared to the physically (and operationally) clear formulations of Clausius and Kelvin, sounds rather abstract and is immediately recognized as the product of a mathematically formed mind. Indeed, Carathéodory was a mathematician well educated in the theory of differential forms and his formulation is clearly rooted in properties of these mathematical objects. Literally it reads:

2TMDL (C. Carathéodory, around 1909, again in Pippard's words):
In the neighbourhood of any equilibrium state of any thermodynamical system there are states inaccessible by any adiathermal process.

The Carathéodory's formulation is obvious in some simple cases like e.g. the adiathermally isolated liquid in which the Joule's paddle-wheel is immersed - adiathermally work can only be added to the system but evidently not extracted from it keeping the volume unchanged and, therefore, states of lower internal energies and the same volume are inaccessible so long as the system remains adiathermally isolated; here however, the Carathéodory's principle is assumed to apply universally to any thermodynamical system, no matter how complicated.

---

[30] It is assumed here that everybody knows what the Carnot cycle or engine is. (Carnot cycles will appears here in Lecture IV). We cannot say that by means of a mechanical device, like the Joule's paddle-wheel one, the work $\bar{W}$ can all be transferred as heat to the reservoir at t_2, for we insisted (in Lecture II) that using paddle-wheel-like devices should be classified as performance of work!

It is fairly straightforward to show that the Carathéodory's principle follows from the Kelvin's principle. The proof is again by *reductio ad absurdum.* Take a thermodynamical system which does not obey the Carathéodory's principle at least in some domain of the manifold of its states. Consider a state A of the system in this domain and let this system make an isothermal change as a result of which it reaches a state B in this domain, absorbing on the way a positive heat Q from a reservoir.[31] The assumption that the Carathéodory's principle is violated in the considered domain means, that all its points are accessible from B on adiathermal paths. So A must be accessible too. If the system is returned to A by an adiathermal path, the total change of its internal energy in the considered transition $A \to$ (isothermal) $\to B \to$ (adiathermal) $\to A$ is zero and therefore, by 1TMDL, the positive heat Q which it absorbed from the reservoir in the isothermal transition from A to B must have been all converted into a positive work $\bar{W} = Q$ in the adiathermal transition. But this contradicts the Kelvin's principle. (For the sake of complete clarity one should admit that it might happen that in some domain the heat form $đQ$ of a system is identically zero. Then any two states within this domain can be connected by an adiathermal path, but vanishing of $đQ$ means that the system is entirely mechanical and not thermodynamical, so in this domain 1TMDL reduces $\Delta U = W$. Such systems indeed do not satisfy the Carathéodory's principle but this does not invalidate, of course, 2TMDL, which applies to thermodynamical systems.)

For completeness one should mention here the Callen's approach to thermodynamics in which 2TMDL takes the form of an axiom in which the existence of entropy as a state function possessing certain properties is postulated and then consequences of this axiom are derived and compared with experimental facts. We will come to discuss this formulation a bit later.

The **mathematical consequences** of 2TMDL are the following

- The differential one-form of heat, $đQ$, of any thermally homogeneous system (that is not possessing internal adiathermal partitions), no matter how complicated,[32] which represents the heat absorbed by the system in a reversible change, is **integrable** (i.e. has an **integrating factor**).

- Among many (mathematically) possible integrating factors of $đQ$ there is one, $1/T$, which is given by a **universal** (i.e. independent of the system) function $T(t)$ of the empirical temperature t, i.e. depending on the system's parameters $X_1, \dots, X_{o-1}, y$ only through the empirical temperature $t(X_1, \dots, X_{o-1}, y)$.

- The exact differential $dS = đQ(X_1, \dots, X_{o-1}, y)/T(t(X_1, \dots, X_{o-1}, y))$ defines, up to a constant, a new state function of the system - its entropy $S(X_1, \dots, X_{o-1}, y)$.

---

[31] If in the isothermal transition from A to B the system loses a positive heat $\bar{Q}$, we can repeat the reasoning exchanging the roles of the states A and B, because isothermal processes can always be made reversible, as I tried to explain in the preceding Lecture.

[32] As explained below, this statement is trivial in the case of simple systems the equilibrium states of which are characterized by two parameters only. But this is not true of the next statement.

- In all adiathermal processes (reversible or irreversible) the entropy does not decrease: $S_{\rm fin} \geq S_{\rm in}$.
- The function $T(t)$ is, up to a scaling factor (setting in fact the units in which it is measured), the absolute temperature as defined by the Carnot cycle.

These consequences can be derived from any of the three formulations of 2TMDL given above. The only difference is the degree of complication of the reasonings involved. Derivation of all these consequences from the Clausius' or Kelvin's formulations of 2TMDL requires, as we will see below, considering "gedanked experiments" invoking cyclic processes (Carnot cycles). The Carathéodory's formulation of 2TMLD seems to be rather abstract but it allows to derive its mathematical consequences without appealing to cyclic processes the use of which (outside the theory of machines producing work or refrigerators) is rather artificial and for this reason it is said to be more economical. However, since in the XXI century we are - or at least we should be! - familiar with rudiments of the theory of forms (I deliberately planned classes so that you get acquainted with them), we can now say that it simply goes in the direction of putting thermodynamics on a postulatory basis, as a purely deductive theoretical system (this tendency, as we will discuss, eventually culminates in the Callen's formulation, celebrated by many, including my respected older colleagues from the 5^{th} floor) for the Carathéodory's principle directly postulates (through the second Carathéodory's theorem) the existence of adiabatic surfaces, that is, it directly states that the heat forms $đQ$ of physical systems do have integrating factor and are therefore integrable, with values of some quantity σ, called empirical entropy, labeling families of the corresponding solutions of the equation $đQ = 0$. But, at least to me, the whole essence of thermodynamics as a phenomenological theory is to demonstrate how these properties of the heat form and the existence of entropy necessarily follow from the formulations of Clausius and Kelvin - the formulations which generalize our direct experience!

As we will be discussing in these lectures, on the practical side 2TMDL

- By the rule $S^{\rm tot}_{\rm fin} \geq S^{\rm tot}_{\rm in}$ determines the direction of real processes occurring in adiathermally isolated (compound) systems.
- Expresses in a precise way and quantifies the degree of irreversibility of physical processes,
- Determines the equilibrium states of thermodynamic systems under various conditions (we will elaborate on this Callenian point of view in due course).
- Puts upper limits on efficiencies of thermodynamic machines or, more generally, limits the efficiency with which heat can be converted into work.
- Limits possible use of huge amount of internal energy of bodies around us (e.g. the energy of oceans).

- Is important for thermodynamics of chemical reactions and in biology and in many other areas (including even economy!).

To quote J.R. Emden:[33] “In the huge manufactory of natural processes the principle of entropy occupies the position of manager, for it dictates the manner and method of the whole business, whilst the principle of energy merely does the bookkeeping, balancing credits and debits.”

---

[33] Jacob Robert Emden (1862 - 1940) a Swiss astrophysicist and metheorologist.

LECTURE IV (TMD)

We now set ourselves to infer the integrability of the heat form $đQ$ of any thermally homogeneous system and, in consequence, the existence of the entropy as a state function, as well as the crucial property of the heat form's integrating factor first from the Kelvin's formulation of 2TMDL. (Since we have shown that this formulation is equivalent to the Clausius' one and vice versa, we may claim, if we wish, that we start from the Clausius' formulation, as well). After this is done, we will discuss how these properties follow from the Carathéodory's principle.

At the beginning we will consider adiathermal and reversible (i.e. adiabatic) changes, in which $đQ = dU - đW = 0$, which form a (rather special) subclass of all adiathermal changes. In the case of simple systems, e.g. fluids characterized by the variables p and V, the equation

$$đQ = \left[\left(\frac{\partial U}{\partial V}\right)_p + p\right]dV + \left(\frac{\partial U}{\partial p}\right)_V dp = 0\,, \tag{42}$$

or, choosing V and $t = t(V, p)$ as independent variables, the equation

$$đQ = \left(\frac{\partial U}{\partial t}\right)_V dt + \left[\left(\frac{\partial U}{\partial V}\right)_t + p\right]dV = 0\,, \tag{43}$$

determines uniquely in the parameter space a family of nonintersecting curves[34] labeled by the initial point which therefore makes the heat form $đQ$ trivially integrable in such cases. If the system is the perfect gas, the equation (43) reduces to

$$C_V^{(t)}\, dt + p(t, V)\, dV = 0\,, \tag{44}$$

and, if in addition the constancy of $C_V^{(t)}$ is assumed, yields immediately, upon using the equation of state $p(t, V) = nRt/V$, the adiabatic curves $t\, V^{nR/C_V^{(t)}} = \text{const.}$, or $p\, V^{1+(nR/C_V^{(t)})} = p\, V^{C_p^{(t)}/C_V^{(t)}} = \text{const.}$ (The relation $C_p^{(t)} = C_V^{(t)} + nR$ satisfied by the principal heat capacities of the perfect gas follows from 1TMDL under the same physical, but as yet mathematically not justified, assumption which allow to write the equation (44).) The heat form $đQ$ is made integrable by just dividing it by t (or const.×t), that is, as advertised, its integrating factor is indeed a function (in this case a linear one) of the empirical temperature t (defined by the perfect gas thermometer).

The example given above is to some extent trivial because in two dimensions every one-form is integrable. (What is less trivial is the fact that the integrating factor is proportional to the empirical temperature.) However in classes you have seen (at least those who attended...) an example (taken from the Pippard's book) of a one-form in three

---

[34] It should be known from the standard Math II course (at least to those trained in math classes by me) that the equations of the type $P(x, y)dx + Q(x, y)dy = 0$ sometimes do have singular points through which more than one integral curve (solution) passes; such cases are here excluded.

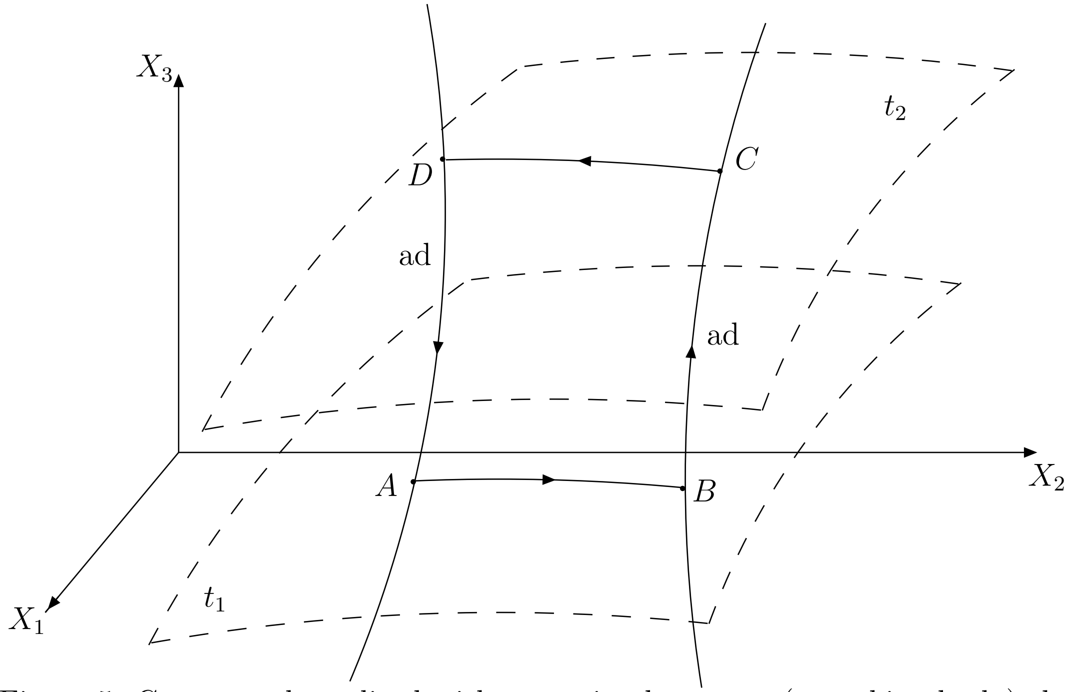

Figure 5: Carnot cycle realized with a nonsimple system (a working body) characterized by three parameters X_1, X_2 and $X_3 \equiv y$. The lines marked "ad" represent adiathermal reversible (adiabatic) changes; the lines AB and CD lie entirely in the two-dimensional surfaces of constant temperature (t_1 and t_2, respectively).

dimensions which is not integrable. Thus not every one-form in more than two dimensions is integrable.

To establish in the general case the integrability of the heat form $đQ = dU - đW$ which can also be written as ($X_{i'}$ stands for all variables other than X_i and for simplicity we have renamed the single nondeformative variable y to X_o)

$$đQ = \sum_{i=1}^{o} \left(\frac{\partial U}{\partial X_i}\right)_{X_{i'}} dX_i - \sum_{i=1}^{o-1} y_i(X_1, \dots, X_o)\, dX_i \,, \tag{45}$$

we will (as in the case of proving the existence of the empirical temperature as a function of state) first present a reasoning which is more physical (and uses virtually no mathematics) and then a more mathematical one.

To show in a physical way, using the Carnot cycles, that the Kelvin's 2TMDL implies integrability of the heat form $đQ$ of a thermally homogeneous system, we start by taking two isothermal hyper-surfaces $t(X_1, \dots, X_o) = t_1$ and $t(X_1, \dots, X_o) = t_2$ with $t_1 \neq t_2$ (if the system has three independent parameters these are ordinary two-dimensional surfaces in the three-dimensional parameter space - see Figure 5) of this system. From 0TMDL we know such hyper-surfaces always exist and are not intersecting with one another (the empirical temperature is a unique state function - it does not have "branches"). We

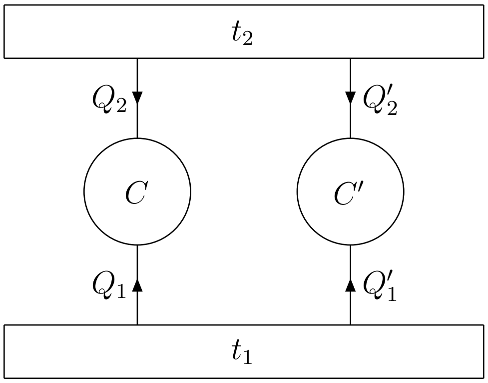

Figure 6: Two different Carnot engines working between the same two heat reservoir's at temperatures t_1 and t_2.

consider on the surface $t = t_1$ a curve connecting two points A and B ($B \neq A$) belonging to this surface. Then we consider two curves which are two solutions of the equation $đQ = 0$ passing through the points A and B and intersecting the surface $t = t_2$ at the points D and C, respectively (Figure 5). (The equation $đQ = 0$ can always be integrated step by step starting from the points A and B and continuing until the resulting two curves cross the hypersurface $t = t_2$, although these curves may not be uniquely determined by the equation $đQ = 0$ and the starting point). In this way one constructs a **reversible cycle** - the Carnot cycle, called also the Carnot engine - the segments AB and CD of which correspond to isothermal changes of the considered system while the segments BC and DA correspond to its adiathermal reversible changes. Beginning from the state A, the system takes in the isothermal change $A \to B$ a heat Q_1 (which can be positive or negative) from a reservoir at the temperature t_1 and in the isothermal change $C \to D$ - a heat Q_2 from another reservoir at the temperature t_2 and remains adiathermally isolated when going from B to C and from D to A. As the system returns to the state A, its final internal energy is the same as the initial one and, by 1TMDL, the work $\bar{W}$ done by the system must be equal

$$\bar{W} = Q_1 + Q_2 \,. \tag{46}$$

Using the Kelvin's statement of 2TMDL it will be now argued that the ratio $-Q_1/Q_2$ is universal, that is, it has the same value irrespectively of the nature of the system performing such a cycle, so long as the two isotherms remain at the (empirical) temperatures t_1 and t_2. The negative sign of the ratio Q_1/Q_2 (that is, the positive sign of $-Q_1/Q_2$) is also a direct consequence of Kelvin's 2TMDL: if it were positive (Q_1 and Q_2 of the same sign), it would be possible to accomplish the cycle (which has been constructed as reversible) in the sense that both Q_1 and Q_2 were positive; of the positive work $\bar{W} = Q_1 + Q_2$ its amount Q_1 could be put back irreversibly (by a Joule's paddle-wheel device, for instance) into the reservoir at t_1 and the net result of the process would be only the extraction of a positive heat Q_2 from the reservoir at t_2 and performance of a positive work $\bar{W} = Q_2$.

This is not possible according to Kelvin. To show that the ratio $-Q_1/Q_2$ is universal, one considers two such reversible Carnot engines which may be constructed using two different thermodynamical systems working between the same two empirical temperatures t_1 and t_2, of which the first engine absorbs the heats Q_1 and Q_2 and the other one Q_1' and Q_2' (Figure 6). It is then possible to chose two integers k and k' so that to the desired accuracy $k|Q_1| = k'|Q_1'|$. This is always possible because any real number $|Q_1|/|Q_2|$ can be approximated by a rational number k'/k (recall the construction of the real number as classes of Cauchy sequences!). One can then treat the two Carnot engines as a single (compound) system and consider its cycle consisting of k runs of the first cycle and k' of the second one accomplished in such senses that the heats Q_1 and Q_1' are of opposite signs. In the complete cycle of the compound system the total heat $k\,Q_1 + k'Q_1'$ taken from the reservoir at t_1 is then zero, while the total heat taken from the reservoir at t_2 equals $k\,Q_2 + k'Q_2'$ and by 1TMDL must be equal to the work done by the compound system. By Kelvin's 2TMDL this cannot be positive, so

$$k\,Q_2 + k'Q_2' \leq 0\,. \tag{47}$$

But because the individual cycles, and therefore the cycle of the compound system, are reversible, also

$$-(k\,Q_2 + k'Q_2') \leq 0\,, \tag{48}$$

which follows from the possibility of accomplishing the cycle of the compound system in the opposite sense. Thus, $k\,Q_2 + k'Q_2' = 0$ too, and from the two equalities

$$\begin{aligned} k\,Q_1 &= -k'Q_1'\,, \\ k\,Q_2 &= -k'Q_2'\,, \end{aligned}$$

it readily follows that $Q_1'/Q_2' = Q_1/Q_2$. Therefore, as proposed, the ratio of heats can only depend on the temperatures t_1 and t_2:

$$-Q_1/Q_2 = f(t_1, t_2)\,, \tag{49}$$

and the function $f(t_1, t_2)$ must be **universal** (independent of the system accomplishing the Carnot cycle between the temperatures t_1 and t_2).

In the next step one shows that the universal function $f(t_1, t_2)$ necessarily factorizes: $f(t_1, t_2) = \phi(t_1)/\phi(t_2)$. To this end one considers a compound Carnot cycle constructed out of one Carnot cycle C_{12} working between t_1 and t_2, consisting as previously of the changes $A \to B \to C \to D \to A$ (Figure 5), and another one, C_{23}, working between t_2 and t_3 and consisting of the changes[35] $D \to C \to F \to G \to D$. Viewing the compound cycle composed of the cycle C_{12} followed by the cycle C_{23} (cf. Figure 7) as a single

---

[35] It is not necessary that these two cycles be performed by the same system; it is enough that the heats $Q_2^{(12)}$ and $Q_2^{(23)}$ taken at t_2 by the systems performing cycles C_{12} and C_{23} were of equal absolute values and of the opposite signs and taken from the same reservoir.

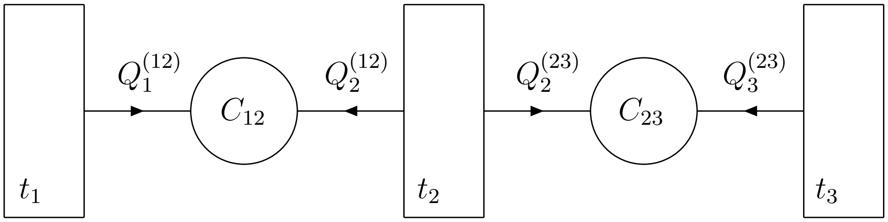

Figure 7: Compound Carnot cycle operating between reservoirs at temperatures t_1 and t_3 composed of two cycles so that the heats taken from the intermediate reservoir at t_2 is zero: $Q_2^{(12)} + Q_1^{(23)} = 0$.

cycle executed between the temperatures t_1 and t_3, on one hand one has $-Q_1/Q_3 \equiv -Q_1^{(12)}/Q_3^{(23)} = f(t_1, t_3)$ while on the other hand

$$-Q_1/Q_3 = f(t_1, t_3) = \left(-\frac{Q_1^{(12)}}{Q_2^{(12)}}\right)\left(-\frac{Q_2^{(23)}}{Q_3^{(23)}}\right) = f(t_1, t_2)\, f(t_2, t_3)\,, \tag{50}$$

because $Q_2^{(12)} = -Q_2^{(23)}$. This is possible only if the function $f(t_1, t_2)$ factorizes as proposed: $f(t_1, t_2) = \phi(t_1)/\phi(t_2)$. Thus, in any Carnot cycle working between temperatures t_1 and t_2, independently of the nature and degree of complication of the system used to construct it,

$$-Q_1/Q_2 = \phi(t_1)/\phi(t_2) \equiv T_1/T_2\,. \tag{51}$$

Obviously $T \propto \phi(t)$ can always be taken for an empirical temperature, but in view of the fact that the ratio of heats taken by any system in the reversible Carnot cycle executed between any two reservoirs is universally given by the ratio of their temperatures T_1 and T_2 defined in this way, T is called the **absolute** or **thermodynamic temperature**. Thus[36]

$$T = \text{const.} \times \phi(t)\,, \tag{52}$$

and operationally the temperature T is determined, assuming that $t_2 > t_1$ and that the cycle is performed so that $\bar{W}$ is positve, by the **efficiency**

$$\eta \equiv \frac{\bar{W}}{Q_2} = \frac{Q_1 + Q_2}{Q_2} = 1 - \frac{T_1}{T_2}\,, \tag{53}$$

of any Carnot cycle working between any two the reservoirs having temperatures t_1 and t_2. This means that any empirical temperature t defined using a thermometric body can

---

[36] As discussed in Lecture I, the proportionality constant in (52) is, since 1954, fixed by assigning to the triple point of water the absolute temperature equal exactly 273.16 K. Before 1954 the constant factor was fixed by the requirement that at 1 atm there be 100 units between the ice melting point and water boiling point; this led to the temperature 271.15 K of the ice melting point (at 1 atm).

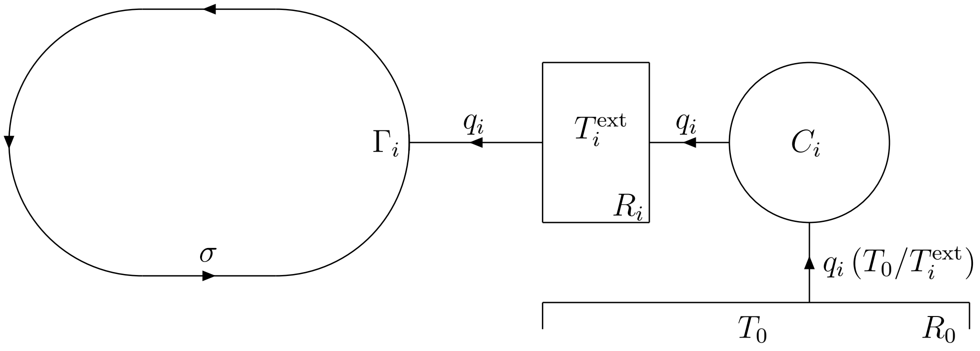

Figure 8: A realization of the fragment Γ_i of the cycle executed by the body σ. The cycle itself, although marked by a continuous line, as if it consisted of a sequence of equlibrium states, need not be such: it need not be reversible and on its stages the system σ may not be in eqiuilibrium.

in principle be calibrated with respect to the absolute temperature, that is the form of the function $\phi(t)$ established, by performing a Carnot cycle using this body as the working substance under conditions as nearly ideal, as possible. In practice, since the perfect gas temperature scale turns out (by taking it as the working substance in the Carnot cycle) to be proportional to the absolute one, it is easier (if the physical conditions allow for this) to relate a given empirical temperature to the perfect gas temperature.

Above we have shown that if a system performs a Carnot cycle between absolute temperatures T_1 and T_2 ($T_1 < T_2$), then

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0\,. \tag{54}$$

We now generalize this result, still relying on the Kelvin's formulation of 2TMDL. The generaliztion is necessery to arrive at the notion of entropy.

Consider a system σ executing a cyclical process of any degree of complexity which we mentally can split into a large number (infinite in the limit) of small (infinitesimal) segments (subprocesses) Γ_i, $i = 1, \ldots, M$ ($M \to \infty$). At the i-th subprocess Γ_i of the cycle a work is done on or by the system σ and some heat is transferred to or abstracted from this system. In the gedanken experiment this transfer of the amount q_i (positive or negative) of heat to the system can be accomplished with the help of an infinitesimal Carnot cycle C_i working between a single reservoir R_0 of temperature T_0 and another auxiliary reservoir R_i at a temperature T_i^{ext} (see Figure 8). One can imagine that the Carnot cycle C_i delivers the amount q_i (positive or negative) of heat from the reservoir at T_0 to R_i and then the same amount of heat q_i is transferred from the reservoir R_i at T_i to the system σ. The last step - the transfer of q_i to σ may be irreversible; we do not assume that the system σ is in equilibrium with R_i nor even in equilibrium in itself, so its temperature (as a single parameter characterizing it) may well not be definable at some or all stages of the cycle.

Since each Carnot cycle C_i is perfectly reversible, it takes, as follows from (54), from R_0 the heat

$$q_i \, \frac{T_0}{T_i^{\rm ext}} \, . \tag{55}$$

Moreover, since all Carnot engines C_i return to their initial states, and the same is assumed about the system σ, the total amount $\bar{W}$ of work obtained from the whole compound cycle equals the total heat taken by all cycles C_i from the reservoir at T_0. By the Kelvin's principle, the work $\bar{W}$ cannot be positive (for this would simply mean taking an amount of heat from R_0 and converting it all into work by means of the cycle executed by the system σ and the auxiliary Carnot cycles). Thus

$$\bar{W} = \sum_i q_i \, \frac{T_0}{T_i^{\rm ext}} \leq 0 \, . \tag{56}$$

Taking the limit $M \to \infty$ and omitting T_0 which is positive, one obtains in this way the inequality

$$\oint \frac{q}{T^{\rm ext}} \leq 0 \, , \tag{57}$$

known as **the Clausius inequality**[37] (despite the fact that we derived it by relying on the Kelvin's principle).[38] The temperature under the integral bears the subscript "ext" to stress (strongly!) the fact that in general this is *not* the temperature of the system σ (this may not be definable) but the temperatures of the (changed in the process) reservoirs from which the heat q is supplied to the system σ and which all are in equilibrium within themselves (they can be assumed to be sufficiently large).

Only if the cycle executed by the system σ is reversible (which requires that at every stage σ is in equilibrium within itself, and in thermal equilibrium with the reservoir R_i with which it exchanges heat at this stage) can one identify the temperatures $T^{\rm ext}$ with the actual (on a given stage of the system's σ cycle) temperature of the system σ (recall - Lecture II - that the heat transfer between two bodies can be realized reversibly only if their temperatures are nearly equal). Moreover, if the cycle executed by σ is reversible, it can be executed in the opposite sense, leading to the inequality

$$- \oint \frac{q}{T^{\rm ext}} = - \oint \frac{đQ}{T} \leq 0 \, . \tag{58}$$

In this case, because the changes of the system σ composing the cycle it executes are reversible, the elementary heat q can be interpreted as the differential form $đQ$ on the space of the system's σ equilibrium parameters (and, as said, $T_{\rm ext}$ identified with the

---

[37] This inequality is frequently written erroneously with the sign $\geq$, because students usually remember that changes of entropy are nonnegative and the factor under the integral $q/T_{\rm ext}$ is confused with $dS = đQ/T$...

[38] In classes you will prove it relying directly on the Clausius's own principle.

actual system's temperature T). Hence, if the cycle is reversible, the Clausius inequality becomes the equality and one is allowed to write it in the form

$$\oint \frac{đQ}{T} = 0\,, \tag{59}$$

which generalizes the equality $Q_1/T_1 + Q_2/T_2 = 0$ holding for Carnot cycles.[39]

The last equality allows to define entropy S as a state function, because it shows that the integral of $đQ/T$ taken between the equilibrium states A and B of any thermodynamical system does not depend on the reversible path (which can be traced in the space of the parameters characterizing equilibrium states of the system) between these states: If $\Gamma^1_{A\to B}$ and $\Gamma^2_{A\to B}$ are two such paths, then

$$0 = \oint \frac{đQ}{T} = \int_{\Gamma^1_{A\to B}} \frac{đQ}{T} + \int_{\Gamma^2_{B\to A}} \frac{đQ}{T}\,, \tag{60}$$

hence

$$\int_{\Gamma^1_{A\to B}} \frac{đQ}{T} = \int_{\Gamma^2_{A\to B}} \frac{đQ}{T}\,. \tag{61}$$

The difference of entropies of the two equilibrium states A and B of a system can be then defined as

$$S_B - S_A = \int_A^B \frac{đQ}{T}\,,$$

with the integral being taken over any **reversible** path connecting A with B, and if for every system a reference state R is chosen and ascribed (arbitrarily) the value S_0 of entropy,[40] then the entropy of any other state A can be defined as

$$S_A = S_0 + \int_R^B \frac{đQ}{T}\,,$$

much in the same way as the internal energy U of every state has been defined with respect to the energy U_0 of a reference state by linking the states by **adiathermal** changes of the system.[41] It is of course reasonable to take the same reference state R of a given system for defining its values of U and S.

---

[39] If in a textbook you see the Clausius inequality written as

$$\oint \frac{đQ}{T} \leq 0\,,$$

you can be sure that the author does not understand it properly.

[40] 3TMDL, to be discussed, allows to fix entropy of any thermodynamic system in absolute terms (up to a common scale factor).

[41] Notice, however, that to define U the adiathermal paths connecting the equilibrium states R and A need not be reversible.

The existence of entropy as a function of state means that the heat form $đQ$ of any thermodynamical system (with the exception of systems possessing internal adiathermal partitions) is integrable - the surfaces which are solutions of the equation (recall the material of classes!) $đQ = 0$ are simply the surfaces $S = \text{const.}$ and its integrating factor, which is singled out by its dependence on the system's parameters $X_1, \dots, X_o$ only through the empirical temperature $t(X_1, \dots, X_o)$ is just the absolute temperature $T = \phi(t(X_1, \dots, X_o))$ defined by the Carnot cycle. In reversible adiathermal changes of any system its entropy S stays constant. This also means that in a reversible processes the heat form $đQ$ can be written as

$$đQ = TdS\,, \tag{62}$$

(and not only as $dU - đW$) and this will have, as will be seen, important practical consequences. Using this relation, which must hold for reversible changes of every system (thermally homogeneous system in the case of a compound one), it is straightforward to show that, granted that the integration factor of the form $đQ$ is restricted to depend only on t (as follows from the presented reasoning), the only its nonuniqueness and nonuniqueness of entropy reduces to

$$\tilde{T} = a\,T\,, \qquad \tilde{S} = \frac{1}{a}\,S + b\,. \tag{63}$$

Indeed, if alternative temperature $\tilde{T}$ and entropy $\tilde{S}$ are introduced so that $đQ$ can be written in two ways as

$$đQ = T\,dS = \tilde{T}(T)\,d\tilde{S}(X_1, \cdots, X_o)\,,$$

(T and $\tilde{T}$ must be functions of t only and this implies that $\tilde{T} = \tilde{T}(T)$), one can pass to the new set of variables (coordinates on the manifold of equilibrium states) with S being one of them, $X_o = S$, and $X_1, \dots, X_{o-1}$ the remaining ones. In these new variables the above equality can be written as

$$dS = \frac{\tilde{T}(T)}{T}\left[\left(\frac{\partial\tilde{S}}{\partial S}\right)_{X_1,\dots,X_{o-1}} dS + \sum_{i=1}^{o-1}\left(\frac{\partial\tilde{S}}{\partial X_i}\right)_{S,X_{i'}} dX_i\right].$$

From this, by comparing the coefficients of the differentials on both sides, it readily follows that $(\partial\tilde{S}/\partial X_i)_{S,X_{i'}} = 0$ (where $X_{i'}$ denotes all parameters other than X_i) which implies that $\tilde{S} = \tilde{S}(S)$ and, moreover, that

$$\frac{\tilde{T}(T)}{T}\,\frac{d\tilde{S}(S)}{dS} = 1\,,$$

or that

$$\frac{d\tilde{S}(S)}{dS} = \frac{T}{\tilde{T}(T)}\,.$$

Since the two sides of the above equality depend on different variables (S and T, respectively) the equality can hold only if $T/\tilde{T}(T) = 1/a = d\tilde{S}/dS$. This gives the result (63).

We can now present another way of arriving at the same consequences of 2TMDL (still relying on the Kelvin's principle) without using Carnot cycles and appealing more to mathematics (but understood physically, without mathematical hieroglyphs). We will illustrate it by considering a three parameter system, so that (45) takes the generic form $đQ = y_1\, dX_1 + y_2\, dX_2 + y_3\, dX_3$. We begin by drawing in the system's parameter space an adiabat, i.e. a line corresponding to an adiathermal reversible change of the system, Γ (constructed by making steps correlated always by the condition $đQ = 0$) crossing two "successive" (i.e. infinitesimally close to one another - they are distributed continuously) isothermal surfaces. Since the system has three parameters, it is possible to draw adiabatic lines beginning at the adiabat Γ and lying entirely within the respective isothermal surfaces. We can therefore consider two points P and P' lying on two such lines on two different but infinitesimally close isothermal surfaces (corresponding to infinitesimally close temperatures t and t') and removed arbitrarily far from the points in which these lines start from the adiabat Γ. The integral

$$\int_P^{P'} đQ\,,$$

taken along the constructed adiabatic lines (the two ones lying on the "successive" isothermal surfaces and the fragment of the adiabat Γ itself) is zero. Then from the Kelvin's principle it follows that if P and P' are infinitesimally close to one another (one can arrange them so, because the two isotherms on which they are situated are infinitesimally close to one another), then the segment PP' must also give $đQ = 0$ (more precisely, the one-form $đQ$ must give zero on the infinitesimal vector PP' joining these two points) for otherwise one would have constructed in this way a reversible cycle in which the system takes the heat $đQ$ on the segment PP' and zero on its remaining parts. Since the cycle would be reversible, this would violate the Kelvin's principle. In this way we can construct a two-dimensional adiabatic surface. Moreover, since we now know that on the vector PP' tangent to this surface the form $đQ$ gives zero, this heat form cannot give zero on a vector PP'' joining P with a neighbouring point P'' not lying in the constructed surface, so PP'' is not an adiabatic line and, therefore, P'' cannot be connected to P by any adiabatic path, however roundabout (if it could, then this path closed with PP'' would represent a reversible cycle wchich could be executed in such a sense that the heat taken by the system on PP'' would be positive; since by 1TMDL this heat would have to be all converted into work, this would violate the Kelvin's principle). So the constructed surface is entirely surrounded by points inaccessible on reversible adiathermal paths. Other adiabatic surfaces can be constructed similarly starting from any other curve along which $đQ = 0$. The reasoning can be generalized to more complicated systems requiring more than three parameters (one constructs in such cases adiabatic hypersurfaces of dimension $o - 1$). The constructed adiabatic (hyper)surfaces can be then labeled by values of a parameter

σ much in the same way as the isothermal (hyper)surfaces were labeled by the values of an empirical temperature t. It is natural to do it in such a way that the resulting function $\sigma = \sigma(X_1, \ldots, X_o)$ is a continuous and differentiable function of the system's parameters.

The existence of adiathermal surfaces can be also demonstrated taking as the starting point the Carathéodory's formulation of 2TMDL, instead of the Kelvin's one. One possible (more mathematical) construction is given in the material prepared for classes. Here we will briefly show the same in a more intuitive way. The Carathéodory's principle states that in the neighbourhood of any state of any thermodynamical system there are other states which are inaccessible in adiathermal changes. From this it follows that there must be even more points inaccessible in reversible adiathermal changes which form a more narrow class than all possible adiathermal changes. It is straightforward to see that if all states of the system are on equal footing, the nearest points inaccessible in reversible adiathermal changes from a given point P must be already infinitesimally close to it (and not at a finite distance from it). For if Q, the nearest inaccessible point on such paths were at a finite distance from P, on the line connecting P with Q, there would have to exist a point P', arbitrarily close to Q, accessible on reversible adiathermal paths from P but not from Q (recall we consider here reversible changes, so any reversible adiathermal change can occur in both directions). Then the nearest to Q point inaccessible from it adiabatically would be infnitesimally close to it (the point P') but not infintesimally close to P - some points would be then distinguished. Hence, the Carathéodory's principle requires that points inaccessible on adiabatic (reversible adiathermal) paths are in fact always infinitesimally close to any equilibrium state of any thermodynamical system.

To show that this implies the integrability of the form $đQ$ of any thermodynamical system not possessing adiathermal internal partitions we first show that if P_1 and P_2 are two equilibrium states of a (thermally homogeneous) system, then either P_2 is accessible from P_1 on adiathermal paths (not necessarily reversible) or it is P_1 which is accessible from P_2 in this way, or both these points are mutually accessible in this way. In other words, we would like to show that the situation that the two points are mutually not accessible adiathermally is physically impossible.[42] To this end let us identify equilibrium points of the system by its internal energy U and $o-1$ deformative parameters $X_1, \ldots, X_{o-1}$ related to the works that can be reversibly done on it.[43] Take the system in the state $P_1 = (X_1^{(1)}, \ldots, X_{o-1}^{(1)}, U^{(1)})$ and perform on it reversibly and adiathermally works (that is, so that $đQ = 0$) until the parameters X_i take the same values as in the state $P_2 = (X_1^{(2)}, \ldots, X_{o-1}^{(2)}, U^{(2)})$. The state $P' = (X_1^{(2)}, \ldots, X_{o-1}^{(2)}, U')$ reached in this way is the same as P_2 only if $U' = U^{(2)}$. In this case P_1 and P_2 are obviously mutually accessible on adiathermal (in this case also reversible) paths. If instead $U^{(2)} > U'$, then - as the experience shows - one can always perform on the adiathermally isolated system

---

[42] This can be treated as the proof of the property of such thermodynamic systems which was used (Lecture II) in ascribing to every system's equilibrium state a value of the internal energy U only if we from the beginning assume (what seems reasonable) that internal energy is a property existing independently of the procedure of ascribing its concrete (relative) values to all system's equilibrium states.

[43] Notice that if the system were purely mechanical, the variable U treated here as its o-th coordinate would be redundant being entirely dependent on the remaining $o-1$ coordinates.

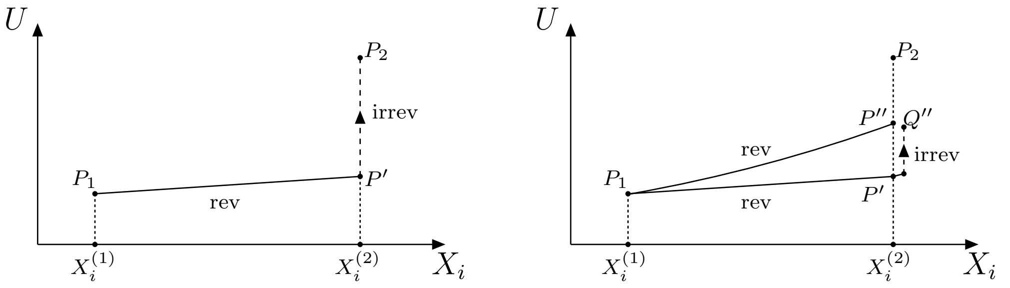

Figure 9: Left: The adiathermal path by which the equilibrium state P_2 is accessible from the equilibrium state P_1. Right: if there were two different (equilibrium) states: P' and P'' with $U'' > U'$ accessible from P_1 on reversible adiathermal paths, every equilibrium state Q'' in the vicinity of P'' would be accessible from P'' on adiathermal paths (in contradiction with the Carathéodory's principle).

an irreversible work (with the help of an analog of the Joule paddle-wheel device or of a current passed through the system) as a result of which its energy increases and reaches the value $U^{(2)}$, while the parameters X_i remain unchanged. In this case P_2 is accessible from P_1 on adiathermal paths (see the left Figure 9). If $U^{(2)} < U'$ one can start from the state P_2, do first on the adiathermally isolated system an irreversible work which increases its energy up to the value U' without affecting the values of the parameters X_i, and then the point P_1 can be reached adiathermally by performing on the system reversible works so that the parameters X_i take the values $X_i^{(1)}$.

Next we show that the point P' reached from P_1 on adiathermal reversible paths (by changing the parameters X_i as a result of works done reversibly on the system) is unique for otherwise the Carathéodory's principle would be violated.[44] Indeed, suppose that from the state P_1 two points $P' = (X_1^{(2)}, \dots, X_{o-1}^{(2)}, U')$ and $P'' = (X_1^{(2)}, \dots, X_{o-1}^{(2)}, U'')$ with $U'' \neq U'$ can be reached by doing reversible works on the adiathermally isolated system. Then if $U'' > U'$, all points in the neighbourhood of P'' could be reached from it on adiathermal paths (contrary to what the Carathéodory's principle says): to reach a point $Q'' = (X_1^{(2)} + \delta X_1, \dots, X_{o-1}^{(2)} + \delta X_{o-1}, U'' + \delta U)$ in the vicinity of P'' one could first reach the state P' from P'' adiathermally and reversibly (passing through P_1); in the same way one could further change the parameters X_i so that they take the values $X_i^{(2)} + \delta X_i$ (the same as the ones defining the point Q''); since the last operation is an infinitesimal change - the increments δX_i being arbitrarily small - the energy of the state reached in this way should be still smaller than $U'' + \delta U$ and the state Q'' could be therefore reached by performing on the adiathermally isolated system an irreversible work not affecting the values of the parameters X_i (see the right Figure 9). If $U'' < U'$, then, of course, all states Q' in the small neighbourhood of the state P' could be reached on adiathermal paths.

Considering now a pair of equilibrium states P_1 and P_2 of a thermally homogeneous system one can ascribe to them values σ_1 and σ_2 of a new parameter σ according to the

---

[44] This follows, of course, already from 1TMDL applied to adiathermally isolated systems which served to define internal energy U (here we consider a more restricted class of works - the reversible ones), but it is instructive to see that the same follows also from the Carathéodory's principle.

rule that $\sigma_2 < \sigma_1$ if P_2 is adiathermally inaccessible from P_1, $\sigma_1 < \sigma_2$ if P_1 is adiathermally inaccessible from P_2 and $\sigma_1 = \sigma_2$ if they are mutually accessible in this way. The procedure is clearly consistent, for if we consider a state P_3 and ascribe to it a value σ_3 by comparing it with the state P_1 (i.e. $\sigma_3 < \sigma_1$ if P_3 is not accessible from P_1 on adiathermal paths, etc.) then if e.g. $\sigma_1 > \sigma_2$ but $\sigma_3 > \sigma_1$, obviously also $\sigma_3 > \sigma_2$ in agreement with the fact P_2 cannot be adiathermally accessible from P_1 through P_3. Furthermore, it is possible to ascribe the values of σ to the equilibrium states of the system in such a way that σ is a single valued, continuous function of the system's parameters $X_1, \dots, X_{o-1}, U$ (and therefore of any other set $X_1, \dots, X_o$ of state parameters uniquely identifying the states of the system). To prove this, it is sufficient to indicate one particular way of ascribing σ satisfying this requirement. This can be done as follows: chose a reference state $P_0 = (X_i^{(0)}, \dots, X_{o-1}^{(0)}, U^{(0)})$ and ascribe to it a value σ_0. To give the value of σ to another point P_1, one can pass from it reversibly (and adiathermally) to a point $P' = (X_i^{(0)}, \dots, X_{o-1}^{(0)}, U')$, which, as has been shown above, is unique. This allows to unambigously ascribe to P_1 the value $\sigma_1 = \sigma_0 + \text{const.} \times (U' - U_0)$, where const.$> 0$ is arbitrary (the physical dimension of σ can be in this way different than that of internal energy).[45] The continuity of σ as a function of the state parameters is then a consequence of the continuity of internal energy U. Furthermore, as follows from the foregoing discussion, if $U' < U_0$, then $\sigma_1 < \sigma_0$ in agreement with the inaccessibility of P_1 on adiathermal paths from P_0. The adiabatic (hyper)surfaces which are solutions of the condition $đQ = 0$ are then defined by the condition $\sigma(X_1, \dots, X_o) = $const.

As the adiabatic (hyper)surfaces can be constructed (the approaches based on Kelvin's and Carathéodory's principles merge at this point), one can now proceed to showing that in a reversible change of a system its heat form $đQ$ can be written as TdS with T being a function of the empirical temperature t only. One starts by labeling the constructed adiabatic (hyper)surfaces of dimension $o - 1$ by a parameter σ (a concrete way of ascribing values of σ, if one starts from the Carathéodory's principle has been given above) constructing thereby a function $\sigma(X_1, \dots, X_o)$, much in the same way as the function $t(X_1, \dots, X_o)$ has been introduced by labeling different isothermal surfaces (and in direct analogy with $t(X_1, \dots, X_o)$ the function $\sigma(X_1, \dots, X_o)$ can be called the **empirical entropy**).

As the adiabatic surfaces (solutions to $đQ = 0$) exist, from the first Carathéodory's theorem[46] it follows that the form $đQ(X_1, \dots, X_o)$ has an integration factor. The same can be also shown here directly by using the method of Lagrange multipliers. Since on surfaces which are solutions to $đQ = 0$, the function $\sigma(X_1, \dots, X_o)$ is constant, in any

---

[45] This way of ascribing the value of σ to a state does not imply that σ is independent of the deformative parameters $X_1, \dots, X_{o-1}$ that characterize that state: if σ is expressed in terms of the internal energy of that state (and not in terms of U' which is energy of another state reached from it adiathermally and reversibly), it will, in general, depend also on the state deformative parameters.

[46] Hopefully discussed in classes.

adiabatic (reversible adiathermal) change of the considered system

$$d\sigma \equiv \sum_{i=1}^{o} \frac{\partial \sigma}{\partial X_i} \, dX_i = 0 \,, \tag{64}$$

similarly as[47]

$$đQ = \sum_{i=1}^{o} y_i \, dX_i = 0 \,. \tag{65}$$

Therefore,

$$đQ - \lambda \, d\sigma \equiv \sum_{i=1}^{o} \left(y_i - \lambda \frac{\partial \sigma}{\partial X_i} \right) dX_i = 0 \,, \tag{66}$$

where $\lambda(X_1, \dots, X_o)$ can be an arbitrary state function (on the space of states of the system). In this relation only $o-1$ differentials dX_i are independent (can be changed at will), because they are (at every point) correlated by (64) or (65). If one regards, dX_1 as determined by the remaining differentials, one can adjust the arbitrary state function $\lambda(X_1, \dots, X_o)$ so as to make the coefficient of dX_1 in (66) vanish:[48]

$$y_1(X_1, \dots, X_o) = \lambda(X_1, \dots, X_o) \, \frac{\partial \sigma}{\partial X_1} \,. \tag{67}$$

Then, since (66) must hold for any adiathermal reversible change in which $dX_2, \dots, dX_o$ can be chosen arbitrarily, one concludes that

$$y_i(X_1, \dots, X_o) = \lambda(X_1, \dots, X_o) \, \frac{\partial \sigma}{\partial X_i} \,, \tag{68}$$

now for all $i = 1, \dots, o$. Hence

$$đQ = \sum_{i=1}^{o} y_i(X_1, \dots, X_o) \, dX_i = \sum_{i=1}^{o} \lambda(X_1, \dots, X_o) \, \frac{\partial \sigma}{\partial X_i} \, dX_i = \lambda \, d\sigma \,. \tag{69}$$

$1/\lambda$ is therefore an integrating factor of $đQ$. One has, however, to show (and this does not follow directly from the Carathéodory's first theorem) that λ cannot be an arbitrary function of the parameters $X_1, \dots, X_o$ but necessarily takes a form of a product of two functions: one which depends on $X_1, \dots, X_o$ through the empirical temperature, and another one which depends on these parameter through the function σ itself.

To this end one considers two systems, 1 and 2, the first one characterized by the variables $X_1, \dots, X_o$ and the second one by $Y_1, \dots, Y_r$, each of which has its own heat

---

[47] Here y_i are not necessarily the generalized forces because we included in them also the terms following from $dU(X_1, \dots, X_o)$.

[48] And this is the essence of the trick with the Lagrange multipliers.

form ($đQ_1$ and $đQ_2$), its own system of labeling adiathermal (hyper)surfaces (functions $\sigma_1(X_1, \ldots, X_o)$ and $\sigma_2(Y_1, \ldots, Y_r)$) and an integration factor of its heat form (functions $\lambda_1(X_1, \ldots, X_o)$ and $\lambda_2(Y_1, \ldots, Y_r)$). We now imagine these two system brought into thermal contact through a diathermal wall and in equilibrium, so that their temperatures are equal

$$t_1(X_1, \ldots, X_o) = t = t_2(Y_1, \ldots, Y_r)\,. \tag{70}$$

The resulting compound system, being thermally homogeneous, is therefore characterized by $o + r - 1$ variables for which one can take the common temperature t and $X_2, \ldots, X_o$ and $Y_2, \ldots, Y_r$. As any thermally homogeneous thermodynamical system, it too must have its empirical entropy function $\Sigma(t, X_2, \ldots, X_o, Y_2, \ldots, Y_r)$ and its heat form $đQ$ must have an integrating factor $\Lambda(t, X_2, \ldots, X_o, Y_2, \ldots, Y_r)$. Since $đQ = đQ_1 + đQ_2$, the following relation must, therefore, hold

$$\begin{aligned}\Lambda(t, X_2, \ldots, X_o, Y_2, \ldots, Y_r)\, d\Sigma(t, X_2, \ldots, X_o, Y_2, \ldots, Y_r)&\\ = \lambda_1(t, X_2, \ldots, X_o)\, d\sigma_1(t, X_2, \ldots, X_o)&\\ + \lambda_2(t, Y_2, \ldots, Y_r)\, d\sigma_2(t, Y_2, \ldots, Y_r)\,.&\end{aligned}$$

Dividing both sides by Λ and going over to the variables $t, \sigma_1, \sigma_2, X_3, \ldots, Y_3, \ldots$ one has

$$d\Sigma(t, \sigma_1, \sigma_2, X_3, \ldots, Y_3, \ldots) = \frac{\lambda_1}{\Lambda}\, d\sigma_1 + \frac{\lambda_2}{\Lambda}\, d\sigma_2\,. \tag{71}$$

It is now clear that Σ is a function of σ_1 and σ_2 only (because only differentials of these two variables appear on the right hand side), so any dependence on $X_3, \ldots, Y_3, \ldots$ and t on the right hand side must drop out. But since λ_1 does not depend on $Y_3, \ldots, Y_r$, neither can Λ, and in the same way, since λ_2 does not depend on $X_3, \ldots, X_o$, the factor Λ cannot depend on these variables either. It then follows, that

$$\lambda_1 = \lambda_1(t, \sigma_1)\,, \qquad \lambda_2 = \lambda_2(t, \sigma_2)\,, \qquad \Lambda = \Lambda(t, \sigma_1, \sigma_2)\,,$$

and, moreover, the dependence on the empirical temperature t must also drop out from the ratios λ_1/Λ and λ_2/Λ:

$$\frac{\partial}{\partial t}\left(\frac{\lambda_1}{\Lambda}\right) = \frac{\partial}{\partial t}\left(\frac{\lambda_2}{\Lambda}\right) = 0\,,$$

and from this it follows that

$$\frac{\partial}{\partial t}\,\ln \lambda_1(t, \sigma_1) = \frac{\partial}{\partial t}\,\ln \lambda_2(t, \sigma_2) = \frac{\partial}{\partial t}\,\ln \Lambda(t, \sigma_1, \sigma_2)\,. \tag{72}$$

All these three derivatives must be therefore equal to the same function of the empirical temperature (the only variable which is common in these derivatives), say $g(t)$, which must be a universal function (the same for all thermodynamical systems). Integrating the first two equalities

$$\frac{\partial}{\partial t}\,\ln\lambda_1(t,\sigma_1)=g(t)\,,\qquad \frac{\partial}{\partial t}\,\ln\lambda_2(t,\sigma_2)=g(t)\,,$$

one gets that

$$\lambda_i(t,\sigma_i)=w_i(\sigma_i)\exp\!\left(\int dt\,g(t)\right),$$

(The functions $w_1(\sigma_1)$ and $w_2(\sigma_2)$ are the "integration constants".) In this way

$$đQ_i=\lambda_i\,d\sigma_i=\left[a\exp\!\left(\int dt\,g(t)\right)\right]\frac{1}{a}\,w_i(\sigma_i)\,d\sigma_i\,, \tag{73}$$

with a a constant, which should be positive (to secure that the absolute temperature - to be identified below - is the increasing function of t) but is otherwise arbitrary.

It remains to show that the integrating factor Λ of the compound system also has this structure, with the same universal function of the empirical temperature (the square bracket in (73)) and the function W which depends on σ_1 and σ_2 only through Σ. From (72), in the same way as above, it follows that

$$\Lambda(t,\sigma_1,\sigma_2)=W(\sigma_1,\sigma_2)\exp\!\left(\int dt\,g(t)\right).$$

Since the factor $\exp\!\left(\int dt\,g(t)\right)$ drops out from the equality $đQ=đQ_1+đQ_2$, one can write (independence of Σ of t follows from (71))

$$W(\sigma_1,\sigma_2)\,d\Sigma(\sigma_1,\sigma_2)=w_1(\sigma_1)\,d\sigma(\sigma_1)+w_2(\sigma_2)\,d\sigma(\sigma_2)\,,$$

from which it is clear that

$$W(\sigma_1,\sigma_2)\,\frac{\partial\Sigma}{\partial\sigma_1}=w_1(\sigma_1)\,,\qquad W(\sigma_1,\sigma_2)\,\frac{\partial\Sigma}{\partial\sigma_2}=w_2(\sigma_2)\,.$$

It is now sufficient to differentiate the first of these two equalities with respect to σ_2 and the second one with respect to σ_1 to arrive at the relation

$$\frac{\partial W}{\partial\sigma_1}\frac{\partial\Sigma}{\partial\sigma_2}-\frac{\partial W}{\partial\sigma_2}\frac{\partial\Sigma}{\partial\sigma_1}\equiv\frac{\partial(W,\Sigma)}{\partial(\sigma_1,\sigma_2)}=0\,,$$

which means that the mapping of (σ_1,σ_2) into (W,Σ) is of rank one, that is, its image is a curve, so that $W(\sigma_1,\sigma_2)=W(\Sigma(\sigma_1,\sigma_2))$. Hence,

$$W(\sigma_1,\sigma_2)\,d\Sigma=W(\Sigma(\sigma_1,\sigma_2))\,d\Sigma\,,$$

which is just what we wanted to prove. Thus the heat form $đQ = đQ_1 + đQ_2$ of the thermally homogeneous compound thermodynamical system can be also written as

$$đQ = \Lambda\, d\Sigma = \left[a \exp\left(\int dt\, g(t)\right)\right] \frac{1}{a} W(\Sigma)\, d\Sigma \equiv T(t)\, dS\,,$$

with T being a universal function of the empirical temperature t and

$$S = \frac{1}{a}\int^{S} d\Sigma\, W(\Sigma)\,,$$

a function which is constant in reversible adiathermal changes. Analogous formulae define the true entropies $S_1 = S_1(\sigma_1)$ and $S_2 = S_2(\sigma_2)$ of its subsystems treated as separate systems. The formulae also clearly show that the freedom in defining entropies of individual (sub)systems and the absolute temperature is that specified in (63). Moreover, as follows from the reasoning, the entropy of a compound thermally homogeneous system (not possessing internal adiathermal partitions) is the sum of the entropies of its subsystems:

$$đQ = đQ_1 + đQ_2 = TdS_1 + TdS_2 = Td(S_1 + S_2) = TdS\,. \tag{74}$$

If a compound system possesses internal adiathermal partitions, or in other words, is composed of subsystems having different temperatures but which are not in thermal contact with one another (are separated by adiathermal walls), no inconsistency arises if its entropy is *defined* as the sum $S = S_1 + S_2 + \ldots$ of entropies of its thermally homogeneous parts. It is clear that defined in this way, entropy of a thermally inhomogeneous compound system stays constant in reversible transitions in which the compound system as a whole is thermally isolated.

In the mathematical arguments used above only reversible adiathermal changes have been exploited. They are sufficient to infer the existence of entropy and proving that the integrating factor of the heat forms of different systems is universal. The Carathéodory's principle itelf makes, however, a statement concerning all possible adiathermal transitions, not only about the reversible ones. In connection with this it should be remarked that it does not specify which points in the vicinity of a given point are not accessible. In particular, it does not tell whether these are poins of smaller or greater entropy than that of a given point. The empirical entropy has been introduced here so that that these are states having (empirical) entropy smaller than the given state and, in the case of changes in which the parameters X_i, $i = 1, \ldots, o-1$ stay fixed, those corresponding to smaller energy U. In fact, it must be said that the Carathéodory's principle must be supplemented with the ancillary law already adopted in assigning energies to states of thermodynamic systems (Lecture II) which asserts that of two states P and P' of a system not possessing internal adiathermal partitions either P' is accessible from P in adiathermal changes or the other way around or both these states are mutually accessible in such changes. Then direct experience (e.g. the Joule experiment) shows that in the case of adiathermal transitions in which deformative parameters of the system stay fixed, accessible are only states of higher energy and we have liberally used this information in defining the empirical entropy. One could proceed somewhat more formally and infer purely mathematically the existence of the empirical entropy from the Carathéodory's principle applied to reversible changes only and then define the true entropy S (along the lines presented above, i.e. exploiting only reversible processes). Then one would have to argue that if the entropies of points P' adiathermally (but not reversibly) accessible from a state P form some range, the entropy S of the point P must be either the lower or the upper limit of this range[49] (in other words, there cannot be points with both, lower and higher entropy, accessible from a given point) for otherwise all points in the neighbourhood of P would be accessible on adiathermal paths from P, contrary to what the Carathéodory's principle states.[50] Further, one would argue that the situation that all states accessible from P_1 have entropy greater than S_1 - the entropy of P_1 - while all states accessible from P_2 have entropy lower than S_2 of P_2 is also impossible.[51] This then would reduce the question whether accessible adiathermally are points of higher or lower entropy to the freedom in chosing the constant a in (63) and (73) thus correlating this question with that about the sign of the absolute temperature T. This will be done below.

Summarizing, the heat form of any thermally homogeneous thermodynamic system is integrable and in reversible changes 1TMDL pertaining to such a system can be written as

$$dU = TdS + \sum_{i=1}^{o-1} y_i \, dX_i \,. \tag{75}$$

This differential relation can be now extended to **all** differential changes, reversible and also irreversible ones, because it is simply the relation between the state functions U, S and X_i at infinitesimally close points (representing equilibrium states of the system). However one should remember, that **only** in reversible changes do the differentials TdS and $y_i dX_i$ have the meaning of the heat absorbed by the system and of the work(s) done on it. In other words, it is only in reversible changes that one can make the identifications

$$q = đQ = TdS \,, \qquad w = đW = \sum_{i=1}^{o-1} y_i \, dX_i \,. \tag{76}$$

To recall the already used example: consider once again a gas expanding into an additional volume of the container when the gas (the whole container) is adiathermally isolated. As

---

[49] One has to reasonably assume that this range is connected.

[50] Taking $X_1, \ldots, X_{o-1}, S$ as state parameters one could accomplish a transition from P changing the value of S in either direction reaching the entropy of a chosen point P' in the neighbourhood of P and then adjust reversibly and adiathermally (so not affecting entropy) all the remaining parameters to the values they have in P'.

[51] If $S_2 > S_1$ then one could reach P_1 adiathermally from P_2 and then in the same way any point in the vicinity of P_2; If $S_2 < S_1$ there would have to exist a state P having entropy between S_2 and S_1 which would be accessible adiathermally neither from P_1 nor from P_2 but then, in agreement with the ancillary law adopted, both these points would be accessible from P, that is P would be a state from which states of greater and lower entropy could be reached adiathermally, which has already been shown to be impossible.

discussed, this change can be accomplished quasistatically, successively opening additional volumes dV (see Figure 3), but is always irreversible. (This can now be quantified by simply comparing entropies of the initial and final states - see Lecture V). In the expansion of the gas into the additional available volume dV certainly $-p\,dV \neq 0$ and, as is clear from 1TMDL also $TdS \neq 0$ because $dU = 0$ (no work is done on the gas and, as the container is assumed to be adiathermally isolated, no heat is absorbed by the system) and the two terms must compensate each other: $p\,dV = TdS$.

The extension we have made here is the basis of the frequently used way of computing the change of a state function (be it the system's energy U or any other its characteristics) in a process $A \to B$ in which the system passes from an equilibrium state A to another equilibrium state B (the process itself may not be reversible) by saying that one can integrate the appropriate forms along any reversible path which connects the states A and B. One does not need to really point this reversible path because all one is doing is just comparing the differential changes of the state functions.

This can be now used to fix the sign of a in (63) and (73). Consider a system adiathermally isolated and perform on it a positive infinitesimal irreversible work not changing its deformative parameters $X_1, \ldots, X_{o-1}$, thereby increasing its energy by $dU > 0$. In view of the relation (13) following from the established (in Lecture II) correlation between the hierarchy of empirical temperatures and hotness, the empirical temperature of the system must have increased by $dt > 0$. The resulting entropy change can be obtained using the trick just stated: one can imagine the same system reaching the new state reversibly through a thermal contact[52] (i.e. now the system is not adiathermally isolated but still the parameters $X_1, \ldots, X_{o-1}$ are kept fixed) with a sequence (infinite in principle) of reservoirs of appropriately adjusted temperatures. Since now no work has been performed on the system, 1TMDL leads to $dU = \text{đ}Q$ and 2TMDL to $dS = dU/T$. If, therefore, the convention is adopted that adiathermally accessible states are those of entropy not lower that that of the actual state, i.e. that by convention $dS > 0$, one is led to the conclusion that $T > 0$ and, therefore, that a in (63) and (73) must be such that $dT/dt > 0$.

Having introduced entropy as a state function we can derive all consequences the phenomenological approach can lead to (see next Lecture). In particular, we can use it to express the heat capacities of systems in reversible processes in a convenient way. Recall first (Lecture II) that any change (process) in which a heat transfer (even an irreversible one due to the fact that the bodies exchanging heat are not at the same temperatures) to or from a system occurs, can be realized quasistatically, i.e. so that the system passes (if other features of the change, like the motion of pistons in the case of fluids, etc., are realized quasistatically so that the work done on the system can be treated as a form on the space of its parameters) through a sequence of consecutive equilibrium states. It is sufficient to imagine that the thermal conductance of the wall through which the system is in thermal contact with the reservoir of heat is very low. This is then equivalent to the

---

[52] This may not always be possible: the example is the system - to be considered in classes - of N isolated spins. Real spins are usually nuclear spins and they are part of a larger system - they can be brought into thermal contact with reservoirs only as part of such a larger system; therefore, if they are treated as isolated, their temperatures can be negative.

trick discussed above being not different from the situation in which heat is transformed reversibly from a sequence of adjusted reservoirs. The system's heat capacity C in such a change can be therefore defined by projecting the system's heat form written as

$$đQ = dU(X_1, \dots X_o) - \sum_{i=1}^{o-1} y_i(X_1, \dots X_o)\, dX_i \,,$$

onto the curve (parametrized by the absolute temperature T) in the space of states representing this change:

$$đQ|_{\text{projected}} = C(X_1(T), \dots, X_o(T))\, dT \,. \tag{77}$$

But if the change is reversible (or is realized as indicated above) one can replace $đQ$ by TdS and write

$$C(X_1(T), \dots, X_o(T)) = T\, \frac{d}{dT}\, S(X_1(T), \dots, X_o(T)) \,. \tag{78}$$

In most cases the process (change) is specified by constancy of certain parameters, e.g. V or p in the case of simple fluids, and their (principal) heat capacities are computed as

$$C_V = T \left(\frac{\partial S}{\partial T}\right)_V \,, \qquad C_p = T \left(\frac{\partial S}{\partial T}\right)_p \,. \tag{79}$$

If the considered system is not simple, more constrains have to be specified to define its heat capacities. For instance, if changes of the volume and the magnetization of a magnetic body are both taken into account, one has to define its heat capacities as, say, $C_{M,V}$ or $C_{M,p}$ etc.

LECTURE V (TMD)

Accepting 2TMDL as a generally valid law of Physics which applies to any macroscopic thermodynamical system, we have shown by a variety of means that the heat form $đQ$ of any (thermally homogeneous) system can, in reversible processes, be written as TdS, where $T = T(t(X_1, \dots, X_o))$ is the absolute temperature (so, the indicator telling us whether two system brought into a thermal contact will be in equilibrium) and $S(X_1, \dots, X_o)$ the new state function - the entropy. In infinitesimal reversible processes 1TMDL, $dU = q + w$, can therefore be written as the sum

$$dU = TdS + \sum_{i=1}^{o-1} y_i \, dX_i \,, \tag{80}$$

of two (inexact) differential forms on the space of the system's equilibrium states, of which the first one represents the heat absorbed by the system in such a process, and the second one is the sum of works done on it. We have in this way obtained a convenient representation of heat capacities characterizing the system in reversible processes as derivatives of its entropy

$$C_X = T \left(\frac{\partial S}{\partial T} \right)_X \,. \tag{81}$$

We have also stressed that the formula (80) remains valid in any infinitesimal change, as relating changes of the state parameters except that if the change is not reversible, the interpretation in terms of heat and works of its individual terms is lost.

Restricting now the attention to the paradigmatic example of a fluid (as a simple system) one immediately notices that the fundamental relation (80), which in this case takes the form

$$dU = TdS - p \, dV \,, \tag{82}$$

implies various, a priori not obvious, relations. For instance, since U is a state function which can be treated as a function of the variables S and V, one obtains from (82) that

$$T = \left(\frac{\partial U}{\partial S} \right)_V \,, \qquad p = - \left(\frac{\partial U}{\partial V} \right)_S \,. \tag{83}$$

Furthermore, its mixed second derivatives must be equal, which gives the relation

$$\left(\frac{\partial T}{\partial V} \right)_S = - \left(\frac{\partial p}{\partial S} \right)_V \,. \tag{84}$$

This is a relation (not obvious a priori) between two measurable coefficients: the one on the right hand side can be obtained experimentally by keeping the volume of the fluid fixed and transferring to it reversibly a heat which is measured; in this way the change ΔS of the fluid's entropy can be determined as $\Delta S = \Delta Q/T$; if this is divided by the measured resulting change Δp of the fluid's pressure, the right hand side is (although not so easily, as is always the case with calorimetric measurements) determined. The coefficient on the left hand side is much easier to measure: one simply measures the change ΔT of the fluid's temperature resulting from a change ΔV of the fluid's volume in the adiathermal (strictly speaking adiabatic) conditions.

The relation (84) and numerous other relations of this sort which can be derived treating U or S (or other state functions mentioned at the end of Lecture II) as a function of different pairs of variables, constitute ones of the most important predictions of phenomenological thermodynamics. From the mathematical point of view all of them follow rather trivially from the existence of entropy as a state function, that is, from the integrability of the heat form $đQ$; from the physical point of view they are highly nontrivial consequences of 2TMDL, the validity of which is not (as we will show on an illustrative example below) a mathematical necessity: 2TMDL is the macroscopic reflection of the intrinsic working of the Nature.

One more, less trivial, example of such relations: consider a nonsimple system, a fluid or a solid under the hydrostatic pressure, which exhibits dielectric properties. To characterize its equilibrium states three parameters, e.g. p, V and $\mathbf{P}$ are needed and the relation (80) expressing 1TMDL as applied to such a system reads ($\mathbf{P} = \int_V d^3\mathbf{x}\,\boldsymbol{\mathcal{P}} \approx V\boldsymbol{\mathcal{P}}$ is the system's total polarization vector and $\boldsymbol{\mathcal{E}}$ is the electric field in which the system is placed)

$$dU = TdS - p\,dV + \boldsymbol{\mathcal{E}}\cdot d\mathbf{P}\,. \tag{85}$$

Instead of considering the internal energy U of this system, one can form another thermodynamic potential (a state function), call it Φ, defined as[53] $\Phi = U - ST + pV - \boldsymbol{\mathcal{E}}\cdot\mathbf{P}$. It is straightforward to see that

$$d\Phi = -SdT + Vdp - \mathbf{P}\cdot d\boldsymbol{\mathcal{E}}\,. \tag{86}$$

In this way Φ is treated as the state function of the independent variables T, p and $\boldsymbol{\mathcal{E}}$, which are all easy to control experimentally. It also follows that

$$\left(\frac{\partial\Phi}{\partial T}\right)_{p,\mathcal{E}} = -S\,, \qquad \left(\frac{\partial\Phi}{\partial p}\right)_{T,\mathcal{E}} = V\,, \qquad \left(\frac{\partial\Phi}{\partial\boldsymbol{\mathcal{E}}}\right)_{T,p} = -\mathbf{P}\,. \tag{87}$$

Equality of the second mixed derivatives of the potential Φ now implies (among others) the relation

$$\left(\frac{\partial\mathbf{P}}{\partial p}\right)_{T,\mathcal{E}} = -\left(\frac{\partial V}{\partial\boldsymbol{\mathcal{E}}}\right)_{T,p}\,. \tag{88}$$

This shows that there must be a nontrivial relation between two seemingly unrelated phenomena: piezoelectricity - polarization of a material as a result of squeezing it - and

---

[53] Φ is the Legendre transform of U. We will tell more about this transformation later.

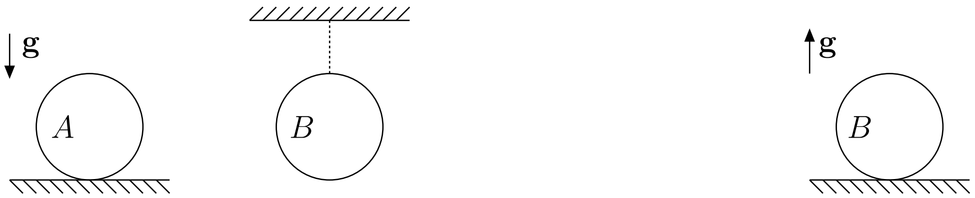

Figure 10: Two identical balls in the Earth's gravitational field. On the right: alternative view on the ball B as placed in the gravitational field pointing upwards.

electrostriction - which is the change of the material's volume when the electric field in which it is placed changes. The two are somehow intimately related by 2TMDL, that is, by the inner working of the Nature.

To illustrate better the deeply **physical** character of 2TMDL, let us consider the following example.[54] Two identical homogeneous balls A and B of mass M made of, say, iron or another material of nonnegligible thermal expansivity, have the same initial temperatures. One of them, A rests on a horizontal plane (in the Earth's gravitational field $\mathbf{g}$), say, on a table, and the other one, B, is suspended on a thread (see the left Figure 10). The same quantities δQ of heat are supplied to both balls. The question one may ask is: which one of the two will then have higher temperature? This problem was once assigned at the International Physics Olympiad and then appeared in many sources. A possible line of reasoning is as follows: the heat supplied to the ball A causes two effects: one is the raising of the ball's center of mass in the gravitational field as a result of the ball's thermal expansion and the other effect is the increase of the ball's internal energy resulting in raising its temperature; so in this case part of the heat δQ is used up for performing a mechanical work. In the case of the ball B instead, thermal expansion lowers the position of its center mass, so in addition to the heat supplied, also the change of its potential energy contributes to increasing its internal energy. The (expected) conclusion, therefore, was that it is the ball B which will have at the end higher temperature. Putting all this in equations: let $\alpha^{\rm lin} = (1/R)(dR/dT)$ be the linear expansion coefficient of the ball and C_0 its heat capacity (which we can take to be independent of the temperature; for simplicity we can assume that all the experiment is carried out at zero external pressure) in the absence of the gravitational field. Then

$$\begin{aligned} \text{ball } A: &\qquad \delta Q = (C_0 + MgR\,\alpha^{\rm lin})\,\delta T_A\,, \\ \text{ball } B: &\qquad \delta Q = (C_0 - MgR\,\alpha^{\rm lin})\,\delta T_B\,, \end{aligned}$$

from which it readily follows, since $\alpha^{\rm lin}$ is clearly positive, that $\delta T_B > \delta T_A$. In the following it will be convenient to treat the ball B as glued to the same horizontal plane as the ball A but in the gravitational field pointing upwards (see the right Figure 10): in this way the two situations A and B are distinguished by the sign of g and one can consider g as varying continuously between positive ($\mathbf{g}$ directed downwards) and negative ($\mathbf{g}$ upwards) values.

---

[54] G. De Palma, M.C. Sormani, *Am. J. Phys.* **83**, 723 (2015).

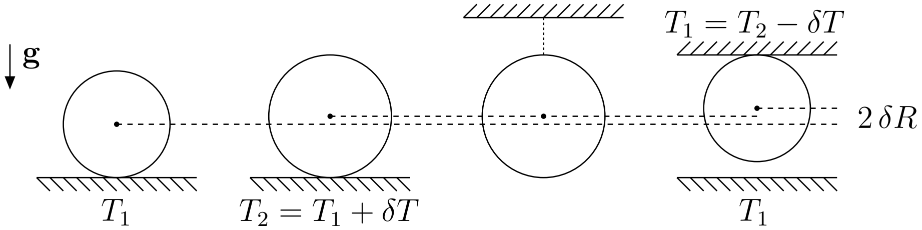

Figure 11: A heat engine exploiting the thermal expansion of the ball.

Using this convention one can introduce the ball's heat capacity at constant gravitational field $C_g(g) \equiv C_0 + MgR\alpha^{\text{lin}}$ as the ball's heat capacity in the gravitational field. Then

$$\left(\frac{\partial C_g(g)}{\partial g}\right)_T = MR\,\alpha^{\text{lin}}\,. \tag{89}$$

Yet this reasonably looking solution has been found (by the authors of the cited paper) to be in conflict with 2TMDL! To see this one can consider the following cycle shown schematically in Figure 11. Start with the ball on the table at a temperature T_1. Then bring it (not changing its position in the gravitational field) into thermal contact with a heat bath of temperature $T_2 = T_1 + \delta T > T_1$. According to the presented reasoning, the ball will absorb heat $\delta Q_{\text{abs}} = (C_0 + MgR\,\alpha^{\text{lin}})\delta T$ and its center of mass will raise by $\delta R = R\,\alpha^{\text{lin}}\delta T$. Now a thread fixed to the ceiling can be attached to it without changing the ball's new position in the gravitational field. Then the ball can be connected to a heat bath at the temperature T_1. According to the "solution" presented above, the ball will loose the heat $\delta Q_{\text{lost}} = (C_0 - MgR\alpha^{\text{lin}})\delta T$ and its center of mass will raise by another $\delta R = R\alpha^{\text{lin}}\delta T$ so that part of the heat absorbed from the reservoir at $T_2 > T_1$ will go into increasing the balls potential energy by $2Mg\,\delta R$. This potential energy can be converted into a mechanical work $\delta\bar{W}$ bringing at the same time the ball to its initial position and completing thereby the cycle (operating between the reservoirs at the temperatures $T_2 = T_1 + \delta T > T_1$ and T_1 with the ball as the working body). One can now ask what is the efficiency η of such a cycle? This is easily computed as the ratio of the work done to the heat absorbed:

$$\eta = \frac{\delta\bar{W}}{\delta Q_{\text{abs}}} = \frac{2MgR\,\alpha^{\text{lin}}\delta T}{(C_0 + MgR\,\alpha^{\text{lin}})\delta T} = \frac{2MgR\,\alpha^{\text{lin}}}{C_0 + MgR\,\alpha^{\text{lin}}}\,.$$

It is clearly independent of the temperature difference δT of the two reservoirs! But according to 2TMDL the efficiency of any cycle operating between temperatures T_1 and $T_2 = T_1 + \delta T$ cannot exceed that of the reversible Carnot cycle

$$\eta_{\text{Carnot}} = 1 - \frac{T_1}{T_2} = 1 - \frac{T_1}{T_1 + \delta T} = \frac{\delta T}{T_1 + \delta T}\,,$$

which decreases to zero as δT goes to zero. Choosing δT sufficiently small, on could, if the presented "solution" were right, beat the efficiency of the Carnot cycle!

What was then wrong? We have assumed that the ball in the gravitational field does not get deformed - that its shape is perfectly spherical independently of whether it rests on the table or is suspended on a thread. This is acceptable as a mathematical assumption but not as a physical one. In reality the ball which rests on the table will be somewhat squashed while the one suspended will be stretched. These deformations do not disappear as $\delta T \to 0$ and ultimately will save 2TMDL. Putting things the other way around: it is 2TMDL which tells us that the deformations cannot be neglected. Microscopically, a solid out of which balls are made is composed of molecules which interact each with the other ones (or at least with the nearest ones). This can be modeled by small masses connected with springs (the simplest model of a solid). The internal energy of the ball is the sum of kinetic energies of the molecules forming it and of potential energies of the springs. In the gravitational fields these springs get either compressed or stretched and this necessarily has some impact on the ball's internal energy independently of the change of the height of its center of mass and the change of the related potential energy which is not included in the ball's internal energy. While a detailed microscopic analysis of all the effects involved would be very complicated, thermodynamics allows to take these effects into account phenomenologically without the need of delving into the microscopic constitution of the ball.

Applying thermodynamics to the ball we must only assume that the ball's internal energy depends on its temperature and on the gravitational field g (here the picture of the ball glued to the table and allowing for the variable sign of g allows to make the analysis simple): $U = U(T, g)$ and that the relevant for the problem parameters T, Y - the ball's center of mass vertical position and g are related by an equation playing the role of the equation of state

$$f(T, Y, g) = 0\,.$$

Applied to the ball 1TMDL takes the form

$$\delta Q = \delta U + Mg\,\delta Y\,,$$

or, if reversible changes are considered, the familiar form

$$dU = TdS - Mg\,dY\,.$$

(The sign of the second term is the same as in 1TMDL $dU = TdS - pdV$ applied to a fluid because in increasing Y the ball must do - if $g > 0$ - a positive work against gravity just in the same way as the fluid must do a work against an external pressure.) The rest is the matter of rudimentary thermodynamical computations: to reduce the necessary steps to the necessary minimum, it is convenient (as always when one controls - here mentally only - a certain parameter like g) to form the analog of enthalpy, defined as $H = U + MgY$, the differential of which is

$$\begin{aligned} dH = TdS + MYdg &= T\left(\frac{\partial S}{\partial T}\right)_g dT + \left[T\left(\frac{\partial S}{\partial g}\right)_T + MY\right] dg \\ &\equiv \left(\frac{\partial H}{\partial T}\right)_g dT + \left(\frac{\partial H}{\partial g}\right)_T dg\,. \end{aligned}$$

Its second form shows that H is treated here as a function of T and g (that is, of the controlled parameters). Since H is a function of state, its second mixed derivatives must be equal:

$$\left(\frac{\partial}{\partial g}\left[T\left(\frac{\partial S}{\partial T}\right)_g\right]\right)_T=\left(\frac{\partial}{\partial T}\left[T\left(\frac{\partial S}{\partial g}\right)_T+MY\right]\right)_g .$$

This leads to the relation (one of the Maxwell identities)

$$\left(\frac{\partial S}{\partial g}\right)_T=-M\left(\frac{\partial Y}{\partial T}\right)_g .$$

Furthermore, since $T(\partial S/\partial T)_g$, as we already know, is just the system's (here ball's) heat capacity at fixed g, the differential dH can be written, using the derived relation, in the form

$$dH=C_g\,dT+M\left[Y-T\left(\frac{\partial Y}{\partial T}\right)_g\right]dg\equiv\left(\frac{\partial H}{\partial T}\right)_g dT+\left(\frac{\partial H}{\partial g}\right)_T dg\,.$$

Applying to this form once again the equality of the mixed second derivatives of H we find that

$$\left(\frac{\partial C_g}{\partial g}\right)_T=-MT\left(\frac{\partial^2 Y}{\partial T^2}\right)_g\equiv-MTY\left[\alpha_{\rm lin}^2+\left(\frac{\partial\alpha_{\rm lin}}{\partial T}\right)_g\right], \tag{90}$$

This is markedly different than the naive formula (89)! In particular, the sign of the derivative is opposite if, as usually happens with real materials, $(\partial\alpha_{\rm lin}/\partial T)_g>0$ (or at least if this derivative is not too large negative to outweigh the positive $\alpha_{\rm lin}^2$ term). Therefore, the heat capacity C of the ball is *larger* if g is negative ($\mathbf{g}$ directed upwards, or when the ball is suspended on a thread) than when g is positive. As a result the correct answer to the problem is $\delta T_A>\delta T_B$.

This example is nice because it clearly illustrates the status of 2TMDL as the physical law. Mathematically one could imagine a world in which the ball is infinitely rigid and its shape does not get deformed when it is placed in the gravitational field. Yet, 2TMDL tells us that in the real physical world this is impossible. 2TMDL, as said, generalizes results of many experiments and by this implicitly takes into account how the real matter behaves and what are the macroscopical consequences of its microscopic (molecular) constitution. The example also illustrates the working of thermodynamics as a phenomenological theory: the internal energy of the ball must somehow be modified by the presence of the gravitational field and although this is (to some extent, at least qualitatively) possible, we do not need to investigate this in detail; we only know that this must be reflected in an equation of state of the form $f(T,Y,g)=0$ which, through 2TMDL, dictates how U and C depend on g; we need, therefore, to find the equation of state and part of the information needed for this is contained in the thermal expansion coefficient $\alpha_{\rm lin}$; measuring

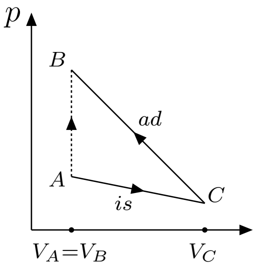

Figure 12: Example of a cycle executed by the perfect gas. The process $A \longrightarrow C \longrightarrow B$ is reversible. The isochoric process $A \longrightarrow B$ can be irreversible or reversible.

this coefficient turns out to be in this case sufficient to tell what will be the answer to the assigned question.

We now return to the Clausius inequality which in Pippard's words holds the clue to the difference between reversible and irreversible processes (changes). It allows to decide whether a given process (occuring in an isolated system) is possible.

Let the two equilibrium states A and B of a system be connected by two paths (processes), $\Gamma^1_{A\to B}$ and $\Gamma^2_{A\to B}$, of which the first path is irreversible, while the second one is reversible. The Clausius inequality (57) applied to the cycle $A \to (1) \to B \to (2) \to A$ can be written in the form

$$\int_{\Gamma^1_{A\to B}} \frac{q}{T_{\rm ext}} + \int_{\Gamma^2_{B\to A}} \frac{q}{T_{\rm ext}} < 0\,.$$

But in the reversible change q can be written as the heat form $đQ$ over the space of the system's parameters and $T_{\rm ext}$ may be identified with the system's temperature T, so

$$-\int_{\Gamma^2_{B\to A}} \frac{q}{T_{\rm ext}} = \int_{\Gamma^2_{A\to B}} \frac{đQ}{T} = S_B - S_A\,,$$

because 2TMDL tells us that $đQ/T = dS$. Thus in the irreversible change $A \to (1) \to B$

$$\int_{\Gamma^1_{A\to B}} \frac{q}{T_{\rm ext}} < \Delta S \equiv S_B - S_A\,. \tag{91}$$

that is, the increase of the system's entropy in any transition (reversible or irreversible) between the states A and B is never smaller than the amount of the "quantity" $q/T_{\rm ext}$ absorbed by the system in this transition. In infinitesimal changes (when the points A and B are infinitesimally close to one another in the space of states)

$$\frac{q}{T_{\rm ext}} \leq dS\,. \tag{92}$$

The Clausius inequality (91) can be simply illustrated by a process in which the perfect gas (for simplicity of constant heat capacity C_V) passes at constant volume V_A

(not necessarily reversibly) from the pressure p_A to the pressure $p_B > p_A$ (see Figure 12). The same final state can be reached reversibly by first expanding the gas isothermally to the volume $V_C > V_A$ and pressure p_C and then compressing it adiabatically (adiathermally and reversibly). Since (as will be obtained in classes) the difference of entropies of the final and initial equilibrium states is $S_B - S_A = C_V \ln(T_B/T_A)$, the inequality (91) applied to the transition $A \to B$ takes the form

$$\int_A^B \frac{q}{T_{\rm ext}} \leq C_V \ln \frac{T_B}{T_A}\,. \tag{93}$$

One can now contemplate different isochoric processes $A \to B$ and the corresponding values of the integral on the left hand side of (93). If the internal energy of the gas is increased by doing on it work using the Joule paddle-wheel device, $q \equiv 0$ and the above inequality is trivially satisfied. The same can be achieved by heating the gas irreversibly by bringing it into thermal contact with a reservoir of heat of constant temperature $T_{\rm ext}$; the heat will flow from the reservoir to the gas provided $T_{\rm ext} \geq \max(T_A, T_B) = T_B$, so

$$\int_A^B \frac{q}{T_{\rm ext}} = \frac{1}{T_{\rm ext}} \int_A^B q = \frac{Q_{A\to B}}{T_{\rm ext}} = \frac{C_V(T_B - T_A)}{T_{\rm ext}} \leq \frac{C_V(T_B - T_A)}{T_B}\,.$$

(In the next to last step we have taken into account that $Q_{A\to B} \equiv U_B - U_A$ is also equal to $W_{C\to B} = C_V(T_B - T_C) = C_V(T_B - T_A)$, as can easily be computed using the perfect gas adiabat equation $p\, V^{C_p/C_V} = \text{const.}$, and the relation $C_p - C_V = nR$, because in the isothermal transition $A \to C$ the energy of the perfect gas does not change). The inequality (93) is then equivalent to the inequality $1 - x \leq -\ln x$ which is true because $0 < x \equiv T_A/T_B < 1$. Finally one can consider a reversible isochoric heating of the gas by connecting it with a sequence of reservoirs of appropriately increasing temperatures; then (using the fact that in the isochoric process $đQ = dU$ and $dU = C_V dT$ in the case of perfect gas)

$$\int_A^B \frac{q}{T_{\rm ext}} = \int_A^B \frac{đQ}{T} = \int_{T_A}^{T_B} \frac{dU(T)}{T} = C_V \int_{T_A}^{T_B} \frac{dT}{T}\,,$$

in which case the inequality (93) is satisfied as equality.

Returning to the general considerations, the very important case arises when the system is adiathermally isolated during the change $A \to B$, which means that $q = 0$. The inequality (91) takes then the form

$$\Delta S \geq 0\,. \tag{94}$$

In the particular case of a (completely) isolated system this is the celebrated **entropy increase law** or just

The entropy law:
The entropy of an (adiathermally) isolated system can never diminish.

Clausius expressed it in the characteristic categorical German (Prussian) way: Energie der Welt ist konstant. Entropie der Welt strebt einem Maximum zu.[55]

The entropy law gives a thermodynamical criterion allowing to decide which processes of those which could occur in adiathermally isolated systems (are allowed by 1TMDL) can actually occur. Following Pippard we are now going to briefly discuss its operation on three simple examples.

Consider first two bodies at different temperatures T_1 and T_2, say $T_2 > T_1$, which together form an isolated system (that is a system with an internal adiathermal partition). If the bodies are brought momentarily into thermal contact, a portion of heat $q > 0$ will flow from the hotter one (that at T_2) to the colder one. As a result the entropy of the hotter one will decrease by[56] $\Delta S_2 = -q/T_2$ and that of the colder one will increase by $\Delta S_1 = +q/T_1$. The total change of entropy of the entire isolated system is positive:

$$\Delta S = \Delta S_2 + \Delta S_1 = q\left(-\frac{1}{T_2} + \frac{1}{T_1}\right) > 0\,.$$

The reverse flow of heat, though consistent with 1TMDL, is forbidden by the entropy law.[57]

As the second example, consider a quantity of a gas in one part of an adiathermally isolated cylinder separated from the second empty part of the cylinder. If the wall separating the two parts is removed, or just a hole is pierced in it, the gas will expand (adiathermally and) irreversibly and its entropy will increase (as you will calculate discussing in classes this so-called Joule process in more detail).

In these two simple examples the changes the discussed systems were undergoing the transitions from one equilibrium state to another equilibrium state as a result of **altering (weakening) the constraints** to which the considered systems were subjected (in the first case the constraint could be represented by an adiathermal wall which then got replaced by an diathermal one, while in the second case the constraint was the wall impermeable to the gas molecules which got removed). In these two cases the operation of the entropy law is clear.

The third example is a moving body which comes to rest due to the presence of friction. In this case the decrease of the body's kinetic energy is accompanied by a

---

[55] This should not be taken too literally: it is disputable whether the (expanding) Universe can be taken as a thermodynamic system and moreover the notion of energy becomes more complicated in general relativity, so we should rather avoid applying thermodynamics to the Universe as a whole.

[56] Stated quickly in this way this may seem to be at odds with the inequality (92). But since the body at T_2 (T_1) which looses (absorbs) a small amount q of heat makes a transision between two infinitesimally close equilibrium states, the change dU of its internal energy must be related by $dU = TdS - p\,dV$ to the changes of its volume and entropy; therefore, if it is assumed that the volume of the body is not changed (or it is so small, that $|p\,dV| \ll |TdS|$), it follows from 1TMDL that $|q| = |dU| = T|dS|$ and the statement is justified.

[57] 2TMDL gives in fact the proof of the possibility to consistently correlate the scale of temperature with the direction of the heat flow (which served to define operationally the notions of “hotter” and “colder”). Logically, this requires formulating 2TMDL without reference to hotter and colder, that is not to base it on the Clausius formulation.

(slight) increase of its and of the surrounding's temperatures. The entropy of the entire system (body and the surrounding) has in this way increased. In this example the initial state is not, strictly speaking, an equilibrium state, but here entropy of the moving body which is in equilibrium within itself in its own rest frame can without any inconsistency be defined as the entropy of the body in its rest system (in relativistic treatment of the body, when $v \lesssim c$, this could be problematic). In general, however, with the exception of trivial situations like that in the third example, the entropy law is (directly) applicable to transitions between equilibrium states only.

Thus, promoting to a valid principle the observation that for a given set of constraints to which it is subjected, an isolated thermodynamic system has only one true equilibrium state, the entropy law can be formulated as the statement (Pippard again)

Entropy law:
It is not possible to vary the (internal) constraints of an isolated system in such a way as to decrease the entropy.

This formulation stresses the role of constraints and will be the basis of the Callen's formulation of thermodynamics which we will discuss shortly. Before that, we will use it (applying it to the second example quoted above) to consider briefly the role of fluctuations. When the gas fills the whole cylinder and is in equilibrium, its local density ρ seems to be uniform. It is however subjected to continuous minute fluctuations most of which are practically undetectable by macroscopic measuring instruments. In reality there is a continuous spectrum of fluctuations ranging with decreasing probability from very small to very large ones. So, very rarely a large fluctuation can occur, e.g. such that - taking things to the extreme - the entire gas spontaneously concentrates in the smaller but macroscopic volume[58] (which it originally left as a result of removing the separating wall or piercing a hole in it). The question may be then asked: what happens to the entropy of the gas during such a large-scale fluctuation? The correct (though perhaps somewhat surprising) answer is: nothing! The thermodynamical entropy which we are considering here[59] stays unchanged. Fluctuations, the spectrum of which is almost continuous - from the minute to the largest possible ones - are part of the thermodynamic equilibrium state and do not represent departures from the equilibrium. The entropy S which is ascribed to an equilibrium state of the system is not ascribed to one (most probable) of its microscopic configurations but to the complete set of microscopic configurations the system

---

[58] Of course the probability of such a fluctuation in a gas consisting of $\sim 10^{23}$ molecules is so fantastically small that there is practically no chance to observe such a fluctuation observind a real system, even waiting as long as the universe's lifetime.

[59] In the kinetic theory of gases one deals with a quantity called H - introduced by Boltzmann - which is usually identified with entropy or, more precisely, with $-S/k_{\rm B}$, but in contrast to the thermodynamic entropy its time evolution can be followed. The famous Boltzmann H-theorem states that H always (although this "always" also requires some qualifications) decreases with time, so $-k_{\rm B}H$ exhibits a property which makes it similar to the thermodynamical entropy. It should be stressed, however, that $-k_{\rm B}H$ should be more properly called the kinetic entropy and it is only in the infinite time limit (and in the thermodynamical limit) that it can legitimately be identified with the thermodynamic entropy (the time evolution of which cannot even be discussed in view of the fact that it is well defined on equilibrium states only).

can be in. This will become clear in the statistical approach which we will discuss later. The important lesson which should be drawn from these considerations is that entropy S (and other state functions) must be regarded as a property **of the system and of its constraints** - in the considered example of the gas and of the cylinder. Only in this way can one understand the statement that S is a function of the gas internal energy U and its volume V (the volume of the cylinder in effect, and not of the volume occupied by the gas at a particular instant). Therefore if the gas is in a smaller volume because of the wall (it is subjected to a stronger constraint) its entropy has one value and it has another value, when it is in the larger volume (weaker constraint) and it is the very act of removing the wall (or piercing the hole in it) which increases the thermodynamical entropy, that is, the act of changing the constraints. It follows that in thermodynamics one never talks about a process of "coming to equilibrium" during which entropy gradually increases. Once the wall preventing the gas to expand into the larger volume has been removed, the entropy increases[60] and the microscopic configuration of the gas (which for a short while still occupies mostly only the initial smaller volume) is now treated as a huge (very improbable to occur spontaneously, as we have said, but here fabricated by an external agent) fluctuation which however is part of the set of all microscopic configurations the gas can assume being subjected to the weakened constraints (i.e. confined in the larger volume).

Similarly in the first example considered, it is the act of replacing the adiathermal wall separating the two bodies by a diathermal one which increases the system's entropy, and not the subsequent flow of heat. Once the diathermal wall is introduced, the system is treated as finding itself in a huge (very improbable to occur spontaneously) fluctuation of the distribution of the kinetic (and potential) energy between all its molecules, but a fluctuation which is part of the new equilibrium state.

Thus any thermodynamic change (a process) should be viewed as a change of the constraints and (one of) the central problem(s) of thermodynamics is, as Callen defines it, to determine the equilibrium state corresponding to the given set of new constraints. It is therefore the second way of formulating the entropy law which is the most adequate one, because it stresses the essential role of constraints to which the system is subjected.

It is interesting to follow here further the discussion of these matters presented by Pippard for it is instructive and sheds light on how thermodynamics as a phenomenological theory lives out of reasonable idealizations. If, says Pippard, one follows this point of view on entropy to its ultimate logical consequences, one should come to the conclusion that the entropy of the Universe is fixed once for ever because no real walls are absolutely impermeable to matter nor no walls are perfectly adiathermal. The state of the Universe we are contemplating should be then viewed only a as huge fluctuation of a more or less uniform density and temperature, that is as a huge fluctuation which is part of an equilibrium state. But, says Pippard, leaving aside the question whether the expanding universe can be treated as an isolated system, such a point of view is not useful and does not allow to make any predictions. A more pragmatic attitude is to make reasonable

---

[60] One is tempted to add "instantaneously" but it is better to avoid notions referring to a time duration.

compromises, that is, to relay on reasonable idealizations: although no walls are absolutely impermeable to matter, on time scales relevant to observed processes one can treat some portions of the Universe as isolated. This is similar to the state of a metastable equilibrium as that of a mixture of oxygen and hydrogen which can for most purposes be treated as a true equilibrium state because the chemical reaction between the two gases, if not artificially stimulated, proceeds at a negligible rate. One is then able to define entropies of physical systems of interest, apply to them the entropy law and make valuable predictions.

It should be, however, noted that the point of view that entropy and other state functions of the system are determined by the constraints, which is natural in that fluctuations find their place in the scheme of thermodynamics, entails a some somewhat strange consequence that the entropy law may seem not to be universally valid: suppose the two bodies at different temperatures are separated by an adiathermal wall. The total entropy of the system is $S_1 + S_2$. Let them contact through an diathermal wall. Then, according to the view adopted above, the entropy of the system instantaneously increases and becomes larger than $S_1 + S_2$, most likely by a significant amount. But if the thermal contact of the bodies is broken before their temperatures equalize, their entropy decreases to nearly the initial one. However, one should firstly remark that in the above reasoning one talks about time which is nonexistent in thermodynamics. Furthermore, one should notice that in the complete experiment the total entropy, nevertheless, does not decrease and the difference of the temperatures of the bodies cannot increase - no useful decrease of entropy can be obtained in this way. Such "paradoxes" can be, therefore, tolerated.

The origin of such "paradoxes", which may cast some doubts on the deduction of the entropy law from 2TMDL, is some inconsistency in viewing large scale fluctuations. Developing the laws of thermodynamics we have adopted the view that no fluctuation can lead to any observable temperature difference of the two bodies in thermal equilibrium - this enabled us to ascribe to them the same temperature. But now we are saying that entropy is determined by the constraints, so the state of two bodies in thermal contact but not yet at the same temperature is treated as a huge fluctuation (which has no chance to be observed if it were to occur spontaneously) which is part of the new equilibrium state. To be consistent one would have to distinguish the temperature of the new equilibrium state (also a function of the constraints) from the imperfectly defined instantaneous temperatures of the individual bodies.

This discussion, says Pippard, leads us into rather deep waters and it is not very useful to continue it within the framework of classical thermodynamics - the proper framework for it being the statistical thermodynamics (in the large meaning of this term - see Lecture I), or even the kinetic theory. But the difficulties just discussed should not be taken as disqualifying the view that entropy is determined by the constraints. The entropy law, although it seems to be violated in the useless way in experiments of the sort mentioned before, is always **valid in practice** and correctly determines what changes are permitted by 2TMDL.

The entropy law has in fact larger range of validity than could be supposed from the foregoing discussion which might be taken to suggest that fluctuations are not accounted by it and that they can in some circumstances lead to its violation. There are in fact strong indications that fluctuations cannot be used to violate the entropy law or 2TMDL in general. It is Maxwell himself who hypothesized a "demon" (called the Maxwell demon ever since) which could control a trapdoor connecting two vessels filled with a gas having the same temperature in both. The demon was supposed to allow to pass from, say, the left vessel to the right one only those molecules which have velocity above the average and in the opposite direction only the molecules of velocity lower than the average. In this way the demon was supposed to be able to rise the temperature of the gas in the right vessel and to lower the temperature of the gas in the left vessel decreasing thereby the total entropy. This is nothing else but an attempt to systematically exploit fluctuations (here fluctuations of the energy of the gas in the region near the trapdoor) to violate 2TMDL. But this way of presenting things assumes that the entropy of the demon itself does not enter the problem nor that it does generate any entropy by its action. Brillouin has analyzed this problem and found this way of reasoning to be unjustified. To distinguish the position and velocity of a molecule (to decide whether to let it pass the trapdoor or not) the demon must be provided with a small flash-lamp (in a gas at uniform temperature T the thermal radiation of the molecules is also uniform and does not allow to distinguish molecules) and the flash-lamp by the radiation it emits operates irreversibly and increases entropy. Brillouin has shown that the decrease of entropy which can be achieved in this way owing to the segregation of the molecules is always overcompensated by the entropy generated by the demon's operation. (This gedanken experiment should be treated at the same footing as the famous Heisenberg gedanken experiment with the microscope, which showed that the quantum mechanical uncertainty principle $\Delta p \Delta q \geq \hbar$ cannot be circumvented and other similar gedanken experiments invented by N. Bohr in the course of his famous discussions with A. Einstein). It is therefore not true, says Pippard, that 2TMDL is only statistically true, being repeatedly violated microscopically (but never seriously, on macroscopically perceptible scales). If entropy is understood as proposed here, as a function of constraints (which as said, allows to incorporate fluctuations into the scheme of phenomenological thermodynamics), 2TMDL is universally valid.

LECTURE VI (TMD)

The view discussed in the preceding Lecture, that it is the constraints, to which a thermodynamic system is subjected, which determine the system's entropy is the basis of the Callenian formulation[61] of thermodynamics which we will discuss now. To some, giving a fourth (after the ones of Clausius, Kelvin and Carathéodory) formulation of 2TMDL may seem superfluous, but this law is the heart of thermodynamics and deserves to be understood from different points of view. The Callenian approach will also serve us to introduce into the play the dependence of thermodynamical state functions on the amount of matter involved, quantified by the number (or numbers, in the case of multicomponent systems) of moles that is to go beyond viewing thermodynamic systems as "black boxes". Moreover, by reducing the complete thermodynamical information about a given system to the knowledge of a single thermodynamical potential (as a function of its natural variables) it puts the necessary order into its characterization. Last but not least, the Callen's formulation of thermodynamics, being directly inspired by the equilibrium statistical physics approach to thermodynamical problems which provides methods of calculating the mentioned thermodynamical potentials, constitutes a direct link between the two parts of this Course.

The Callen's formulation of thermodynamics is based as all previous ones on the postulate that there exist equilibrium states, on 0TMDL which allows to introduce an empirical temperature, and on 1TMDL which now will be written in the form $\Delta U = Q + W + Z$ (or $dU = q + w + z$ if infinitesimal changes are considered) into which also the work Z (or z) related to a change of the amount of matter in the considered system has been included. In reversible processes it will be, of course, possible to write as in the conventional approach $w = đW$ and $q = đQ$. If the matter transfer is done reversibly (the necessary conditions for this will become clear in due course) it will be possible to write $z = đZ = \sum_{i=1}^{r} \mu_i \, dn_i$, where μ_i is the chemical potential associated with the i-th material component of the system.

2TMDL in this approach takes the form of the following postulates

- There exists entropy S which is defined on all equilibrium states of any thermodynamic system which is a function of the system's internal energy U and other globally defined deformative parameters like its volume V, magnetization $\mathbf{M}$, and the like (denoted below collectively $X_1, \dots, X_{o-1}$) and, if the system is composed of matter and possesses the property of extensiveness, of the amount of matter contained in the system and represented by the number n (numbers n_j, $j = 1, \dots, r$ in the case of multicomponents systems) of moles of its constituents. If the system is homogeneous and possesses the property of extensiveness, entropy is a homogeneous function of order one[62] of its global arguments.

---

[61] In fact the approach we call Callenian comes directly from Gibbs and was subsequently developed by Tisza; it has been written up by Callen in his textbook on thermodynamics.

[62] A function $f = f(X_1, \dots, X_n, y_1, \dots, y_k)$ is said to be homogeneous of order p in its first n arguments if $f(\lambda X_1, \dots, \lambda X_n, y_1, \dots, y_k) = \lambda^p f(X_1, \dots, X_n, y_1, \dots, y_k)$.

- S is a (differentiable, at least twice) monotonic function of the internal energy U

$$\left(\frac{\partial S}{\partial U}\right)_{X_i,n_j} > 0\,,$$

- Entropy of a system consisting of several subsystems is additive
- Entropy of a system the global parametres U, $X_1,\dots,X_{o-1}$, $n_1,\dots,n_r$, of which are fixed takes on the maximal value with respect to all possible equilibrium states which could be realized with the help of constraints stronger (or not weaker) than the ones the system is actually subjected to.

We will now explain these postulates.

In most cases we deal with systems which are homogeneous or consist of several subsystems which are homogeneous and posses the property of extensiveness which means that their global parameters $X_1,\dots,X_{o-1}$ (those which characterize the system as a whole, e.g. its volume, internal energy, total polarization, etc., in contrast to those which although uniform throughout the whole system when it is in equilibrium, could in principle locally take different values) and the amount of matter contained in them (quantified by the number(s) of moles $n_1,\dots,n_r$) scale, when the system is isolated, proportionally to their internal energy U: if $U \to \lambda U$, then $X_i \to \lambda X_i$ and $n_i \to \lambda n_i$. This amounts to the assumption that if such a system is mentally divided into two or more parts (which are then treated as separate thermodynamical (sub)systems), its internal energy U is the sum $U = U_1 + U_2$ of the internal energies U_1 and U_2 of these two (or more) parts taken separately (separated by a wall) which in turn means that the energy U_{12} of the interaction of these two parts is negligible compared to U_1 and U_2 and also that in the total internal energy surface effects are negligibly small. This of course requires that the intermolecular forces be (effectively) short range.[63] In many cases (e.g. fluids) scaling of the dimensions of the system corresponding to scaling its energy can be reasonably arbitrary, and can change the system's shape. This is not so in the case e.g. of magnetic (or dielectric) systems placed in an external magnetic (electric) field, because the precise form of the magnetic (electric) field inside a magnetic (dielectric) specimen depends on its shape and so does the total magnetization (polarization); however owing to the scale (in fact even conformal) invariance of (free) classical electrodynamics, magnetic and dielectric systems can be treated as extensive with respect to scalings of their size which do not change their shapes. Global parameters U, X_i, n_j of such systems are called **extensive parameters**. One should also remark that thermodynamics can be applied to systems which are not characterized by the numbers of moles, yet are extensive. The most prominent example of such a system is the electromagnetic field in equilibrium with the walls of a cavity which is characterized only by the internal energy U and the volume V of the cavity. Other

---

[63] Although electromagnetic forces are long-range, they always get screened in electrically neutral material systems. The gravitational interactions cannot be screened in this way and therefore systems in which they play important role are not extensive (see the black hole example below).

systems of this kind (of which the electromagnetic field is the simplest example) are relativistic quantum fields which, although in popular accounts identified with different sorts of elementary particles (the misleading and infamous "field-particle duality" which should be definitely banished[64] forever from serious treatments of the quantum field theory!) are nevertheless systems of fields and the numbers of different kinds of particles contained in such a system are from the point of view of thermodynamics and equilibrium statistical physics ill defined quantities. For this reason chemical potential cannot be directly associated with them.[65]

As far as systems possessing the extensiveness property are concerned, it is in many situations convenient to work not with extensive quantities U, $X_1, \dots, X_{o-1}$, like V or $\mathbf{M}$, etc., and $n_1, \dots, n_r$, but with molar quantities u, v, $\mathbf{m}$, etc., and the molar fractions x_i defined as

$$u = U/n\,, \qquad v = V/n\,, \qquad x_i = n_i/n\,, \qquad n \equiv \sum_{i=1}^{r} n_i\,.$$

Internal energy (and any of the state functions, like the already introduced enthalpy H, Helmholtz free energy F or Gibbs function G) of an extensive system which all are necessarily extensive quantities, when written as functions of extensive and intensive parameters like pressure p, can be written in the form

$$U(p, V, n_1, \dots, n_r) = n\, u(p, v, x_1, \dots, x_r)\,,$$

etc. This works also the other way around: if a state function of an extensive system is given for one mole of it, say the molar Helmholtz function $f = f(T, v, x_1, \dots, x_r)$, its full dependence on the number n of moles can always be restored by writing

$$F(T, V, n_1, \dots, n_r) = n\, f(T,\, V/n,\, n_1/n, \dots, n_r/n)\,.$$

Internal energy and Helmholtz free energy of the electromagnetic field must in turn take the forms $U(T, V) = V u(T)$ and $F(T, V) = V f(T)$, respectively.

Entropy of a (sub)system which possesses the property of extensiveness is (and this is part of the postulate) a homogeneous function of first degree of the extensive (global) parameters:

$$S(\lambda U,\, \lambda X_j,\, \lambda n_i) = \lambda\, S(U, X_j, n_i)\,. \tag{95}$$

which means that it too can be written as

$$S(U, V, \dots, n_i) = n\, S(U/n,\, V/n, \dots, n_i/n) = n\, s(u, v, \dots, x_i)\,. \tag{96}$$

---

[64] If there is any meaning to be attached to this notion, it has rather to do with the astonishing fact that quantum states of most systems of relativistic fields at all exhibit features normally associated with particles; however the simplistic identification of particles with a concrete type of field is fundamentally wrong.

[65] Laws governing the dynamics of system of relativistic fields frequently give rise to conserved charges - like the electric charge in quantum electrodynamics - and it is with these charges that in the statistical physics approach chemical potentials must be associated.

One should be aware, however, that not all systems to which thermodynamics applies are extensive. One important example of a nonextensive system is the black hole to which we will devote below a digression. Of course, extensiveness (or its lack) plays no role when thermodynamic systems are treated as "black boxes" without inquiring into their internal material constitution.

The maximum entropy postulate, central to the whole Callenian approach, can be best elucidated on a simple system of a gas enclosed in a container. The total volume V of the container is fixed, similarly as the total gas energy U and the number of its moles n. By introducing different kinds of auxiliary walls internal with respect to the system: movable or nonmovable, adiathermal or diathermal, permeable or impermeable to molecules, it is possible to divide the container into an arbitrary number of cells of different (macroscopic) sizes V_i in which different numbers of moles n_i and portions U_i of energy are blocked in equilibrium states, the only constraints being

$$\sum_i n_i = n\,, \qquad \sum_i U_i = U\,, \qquad \sum_i V_i = V\,.$$

One can imagine in principle infinitely many ways[66] of blocking in this way various equilibrium states of the system (which becomes in this way composed of several subsystems). In agreement with the first postulate of the Callenian approach, each such equilibrium state is ascribed a certain entropy S which depends on U, V and n and on the distributions U_i, V_i, n_i of these quantities between the subsystems into which the system has been divided by the auxiliary walls. It is given by the sum of entropies of the subsystems into which the system has been split. The equilibrium state of the gas in the absence of auxiliary internal walls is that state among all equilibrium states which can be "fabricated" by introducing these walls in different ways, that has the greatest entropy. This Callenian principle of maximum entropy determines in particular the equilibrium states of isolated systems the global parameters of which U, $X_1, \dots, X_{o-1}$, $n_1, \dots, n_r$ by definition have fixed values. It will be later generalized to apply also to nonisolated systems which are in different specific contacts (thermal, mechanical, material) with their surroundings.

One should note a similarity of this principle to the known principle of least action in Mechanics - the trajectory $q^i(t)$ which the system goes from its position $q^i(t_1)$ at t_1 to its position $q^i(t_2)$ at t_2 is that trajectory which gives the smallest value of the action functional $I[q(t)]$. Operationally, one can consider different trajectories $q^i(t)$ with fixed ends, compute the value of the action on each of them and compare these values: the true trajectory is the one giving the smallest value of I. Because of this similarity, the Callenian formulation of 2TMDL is sometimes called variational. Actually, in mechanics the true trajectory may not be the one corresponding to the smallest value of I: it must only be the stationary one.[67] This is not so in thermodynamics: the true equilibrium

---

[66] Of course, one should not consider divisions of the system into subsystems so small, that surface effects could become important.

[67] The reason for this, as well as the justification of the variational principle on which mechanics can be based is, of course, quantum mechanical.

state of the system maximizes its entropy on the set of all **virtual** equilibrium states which can be fabricated by imposing on it arbitrary auxiliary constraints stronger (not weaker) than the ones the system is actually subjected to. One should also notice that in contrast to mechanics, in which the action I has in each case a form known from the beginning, in the variational formulation of 2TMDL the formula for entropy is not given. Only its existence is postulated.[68] Another remark is that it is now clear that with respect to isolated systems, the entropy S plays the role of the thermodynamical analog of the potential (hence, it is one of thermodynamical potentials) as its maximum (over all possible virtual equilibrium states) determines the equilibrium state of any isolated system

It will be noted that the entropy law, to which one arrives only through a combination of arguments, if the Clausius/Kelvin's or Carathéodory's formulations of 2TMDL are adopted, in the Callenian formulation is built in it from the outset: if the internal constrains to which an isolated system is subjected are weakened, the entropy of the new equilibrium state can only be greater than in the old equilibrium state (or at most equal to it) because a larger set of virtual equilibrium states becomes possible.[69] In an isolated system reversible can only be processes (changes of constraints) which happen not to increase the entropy.

The explicit form of the expression

$$S = S(U, X_1, \dots, X_{o-1}, n_1, \dots, n_r)\,, \tag{97}$$

($S = S(U, V, n)$ in the case of a one-component fluid) for entropy of a system characterized by the displayed parameters is called the **fundamental relation in the entropy representation**. As will become clear, if it is known as a function of these global (extensive) variables, our thermodynamical information about the system is complete (this is the meaning of the word "fundamental"). The derivatives of the entropy (of a simple, possibly multicomponent fluid for definiteness) are

$$\left(\frac{\partial S}{\partial U}\right)_{V,n_i} \equiv \frac{1}{T}\,, \qquad \left(\frac{\partial S}{\partial V}\right)_{U,n_i} \equiv \frac{p}{T}\,, \qquad \left(\frac{\partial S}{\partial n_i}\right)_{U,V,n_i'} \equiv -\frac{\mu_i}{T}\,. \tag{98}$$

Although this is suggestive (and obvious for those who already know 2TMDL), in the Callenian approach one has yet to demonstrate that the quantities T, p and μ_i (and

---

[68] The question of uniqness of entropy then arises. Some claim that to eliminate the nonuniqueness of the form $S' = S + \text{Const.} \times U$ yet one more postulate (the Guggenheim postulate) that entropy stays constant in reversible adiathermal changes must be adopted. Within phenomenological thermodynamics this problem does not makes its appearence in practice for the form of S must anyway be reconstructed from the empirical data according to the standard procedures. In turn statistical mechanics gives the concrete, unambiguous prescription how to obtain entropy of a system, if its microscopic dynamics (classical or quantum) is known and in this way leaves no room for such an ambiguity.

[69] It is clear that if the domain of the arguments over which one seeks the maximum (minimum) of a function (functional) is enlarged, the maximum (minimum) can only increase (decrease).

possibly others, if the system is more complicated) **defined** in this way do indeed have the meanings of the ordinary (absolute) temperature,[70] pressure and chemical potentials.

To show that the parameter T defined by the derivative of entropy with respect to the internal energy has indeed the meaning of the possible thermal equilibrium indicator, we consider a system the energy U and the number of moles of which are fixed, say a fluid in a cylinder of fixed volume V. Introducing an adiathermal nonmovable wall which divides the cylinder into two parts of fixed volumes V_1 and V_2 containing n_1 and n_2 moles (one can assume that $V_1/n_1 = V_2/n_2$ so that the gas densities in the two parts are equal; of course $V_1 + V_2 = V$ and $n_1 + n_2 = n$) we can fabricate different virtual equilibrium states corresponding to different distributions U_1 and U_2 of the total energy U between the two subsystems into which the original single system has been split. Alternatively, one can consider a system which is from the beginning composed of two subsystems, two (possibly different) gases in two containers, of volumes V_1 and V_2 and numbers n_1 and n_2 moles or a gas and a solid or even two solids, which are in thermal contact through a diathermall wall; the virtual equilibrium states blocked by replacing the diathermal wall by an adiathermal one correspond to different distributions of the total energy U between the two gases. According to the Callenian entropy maximum principle, the equilibrium state in the complete absence of the wall in the first case or when the wall is diathermal in the second case, is that one of the virtual equilibrium states that has the greatest entropy. Now, entropies of the virtual equilibrium states when the system consists of two subsystems are given (relying on the postulated additivity of entropy) by the formula

$$S = S_1(U_1^{\rm eq} + \delta U,\, V_1,\, n_1) + S_2(U_2^{\rm eq} - \delta U,\, V_2,\, n_2)\,,$$

in which the departures of the virtual state energies U_1 and U_2 from the (unknown yet) equilibrium distribution $U_1^{\rm eq}$, $U_2^{\rm eq}$ ($U_1^{\rm eq} + U_2^{\rm eq} = U$) have been parametrized with δU (automatically taking into account the condition $U_1 + U_2 = U$). Actually, in the example of a gas in a single container, since this system is extensive, the two functions S_1 and S_2 are simply the same function $S(\cdot, \cdot, \cdot)$. The extremum condition $\delta S = 0$ which, in view of the simplified way of seeking the maximum[71] could simply be reduced to $dS/d(\delta U) = 0$,

---

[70] Note that the requirement (which is one of the postulates) that $(\partial S/\partial U)_X > 0$ ensures positivity of the absolute temperature defined here. In the statistical physics part of this course we will consider a system (which in reality must be a subsystem of a larger system) not satisfying this requirement and therefore capable of assuming negative temperatures; we will also briefly discuss how 2TMDL should be generalized to apply to such systems.

[71] This reasoning in which one considers only a certain rather narrow class of all possible virtual equilibrium states which could be fabricated should be confronted with the "practical" approach to the determination of the trajectory $q^i(t)$ of a classical particle: assuming we do not know the variational calculus (which reduces the problem to solving the Euler-Lagrange differential equations) we invent a trial trajectory which connects the initial and final system's positions in the time interval $t_2 - t_1$ and depends on one (or a few) free parameter(s) λ (λ_i). The action computed on this trial trajectory depends therefore on this (these) parameter(s) and can be minimized with respect to it (them). Of course, finding the true trajectory in this way is possible only if it is one of the trial trajectories taken into account (which form only a very narrow class of all possible virtual trajectories).

gives the equality

$$\delta S = \left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,n_1}\Bigg|_{U_1=U_1^{\rm eq}} \delta U + \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,n_2}\Bigg|_{U_2=U_2^{\rm eq}} (-\delta U) = 0\,.$$

which implies that in equilibrium the parameters $1/T_i = (\partial S_i/\partial U_i)_{V_i,n_i}|_{U_i=U_i^{\rm eq}}$, $i = 1, 2$, must be equal.[72] Thus the derivative of S with respect to U plays the role of the thermal equilibrium indicator i.e. of an empirical temperature. Defining it to be $1/T$ and not T leads to the assignment of higher temperatures to hotter bodies. Indeed, consider one of the virtual equilibrium states with $\delta U \neq 0$ but infinitesimal $|\delta U| \ll U$ as a **real** state of **two bodies** and assume that, say,

$$\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,n_1}\Bigg|_{U_1=U_1^{\rm eq}+\delta U} < \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,n_2}\Bigg|_{U_2=U_2^{\rm eq}-\delta U} ,$$

that is, as follows from (98), that $T_1 > T_2$. Then, when the adiathermal wall is replaced by an diathermal one, the system will reach the new equilibrium state and in reaching it the two bodies will exchange energy (in the form of heat) in such a way that the entropy will increase:

$$\Delta S = \Delta S_1 + \Delta S_2 = -\frac{\delta U}{T_1} + \frac{\delta U}{T_2} = \left(\frac{1}{T_2} - \frac{1}{T_1}\right)\delta U > 0\,.$$

This means that δU must be positive, i.e. that the body which had higher temperature looses energy and that of lower temperature gains it.

Similarly, to show that the parameter p defined by the second one of the derivatives (98) is the ordinary pressure, we consider virtual equilibrium states of the system (again, for concreteness let it be a gas in the cylinder) which can be realized with the help of the rigid adiathermal wall but now without correlating the mole numbers n_i with the volumes V_i (that is, allowing for different densities of the gas in the two parts) and allowing for different distributions of the internal energies U_1 and U_2 and of the volumes V_1 and V_2. Again one can alternatively consider two (possibly different) gases, n_1 and n_2 moles of each, in two parts of a cylinder separated by an unmovable adiathermal wall. The values $U_i^{\rm eq}$ and $V_i^{\rm eq}$ which will be realized in the equilibrium state which the system assumes in the absence of the wall in the first case and if the wall is movable and diathermal in the second case, are the same as in that virtual state obtained in the presence of the wall, which maximizes the total entropy

$$S = S_1(U_1^{\rm eq} + \delta U,\, V_1^{\rm eq} + \delta V,\, n_1) + S_2(U_2^{\rm eq} - \delta U,\, V_2^{\rm eq} - \delta V,\, n_2)\,,$$

---

[72] The assumption - which is part of the Callenian postulates - that entropy is a monotonic function of the internal energy automatically means that at $U_i = U_i^{\rm eq}$ the first derivative $dS/d(\delta U)$ changes sign, so this point is the extremum.

that is such that

$$\delta S = \left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,n_1}\Bigg|_{U_1=U_1^{\rm eq},V_1=V_1^{\rm eq}} \delta U + \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,n_2}\Bigg|_{U_2=U_2^{\rm eq},V_2=V_2^{\rm eq}} (-\delta U)$$
$$+ \left(\frac{\partial S_1}{\partial V_1}\right)_{U_1,n_1}\Bigg|_{U_1=U_1^{\rm eq},V_1=V_1^{\rm eq}} \delta V + \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2,n_2}\Bigg|_{U_2=U_2^{\rm eq},V_2=V_2^{\rm eq}} (-\delta V) = 0\,. \qquad (99)$$

Since the departures δU and δV can be varied independently (in fabricating different virtual equilibrium states), this implies equality (in the equilibrium state realized in the absence of the wall) of the two temperatures $1/T_i = (\partial S_i/\partial U_i)_{V_i,n_i}|_{U_i=U_i^{\rm eq},V_i=V_i^{\rm eq}}$ and of the two pressures $p_i = T_i(\partial S_i/\partial V_i)_{U_i,n_i}|_{U_i=U_i^{\rm eq},V_i=V_i^{\rm eq}}$. The reason, why the derivative of S with respect to V is identified with p/T and not with p, can be seeked in the "shocking" relation

$$\left(\frac{\partial S}{\partial V}\right)_{U,n} = -\left(\frac{\partial S}{\partial U}\right)_{V,n}\left(\frac{\partial U}{\partial V}\right)_{S,n} = -\frac{1}{T}\left(\frac{\partial U}{\partial V}\right)_{S,n}\,.$$

To corroborate the interpretation of the derivative of S with respect to V as the ratio of the system's pressure to temperature, one can also consider n moles of a gas in an adiathermally isolated cylinder fitted with the piston (of cross section area A) on a spring satisfying the Hooke's law (see Figure 13). The total energy E of the entire system (the gas, the piston and the spring) is fixed. Blocking the piston in different positions (characterized by the variable x) one can fabricate different virtual equilibrium states of the system. The equilibrium state assumed by the system when the piston is not blocked, is that one of the virtual states which has the maximal entropy (the entropy of the mechanical elements of the system - the spring - is assumed to be zero; V_0 is the volume of the cylinder when the spring has its free length)

$$S = S_{\rm gas}(E - \frac{1}{2}kx^2,\, V_0 + Ax, n)\,.$$

Equating to zero the derivative of S with respect to x one gets the condition (the symbol $x^{\rm eq}$ stands for $U = E - \frac{1}{2}kx_{\rm eq}^2$ and $V = V_0 + Ax_{\rm eq}$)

$$-kx\left(\frac{\partial S}{\partial U}\right)_{V,n}\Bigg|_{x^{\rm eq}} + A\left(\frac{\partial S}{\partial V}\right)_{U,n}\Bigg|_{x^{\rm eq}} = 0\,.$$

Since kx/A is the mechanical pressure $p_{\rm ext}$ exerted on the gas by the piston, one learns that in equilibrium state $T_{\rm eq}(\partial S/\partial V)_{U,n}|_{x_{\rm eq}} = p_{\rm ext}$.

It is appropriate to comment in this place (also because this sheds some light on the preceding example) on the so-called problem of the adiabatic piston by which one means the problem of establishing the conditions which should determine the equilibrium state of the system consisting of two (possibly different) gases, n_1 and n_2 moles of each,

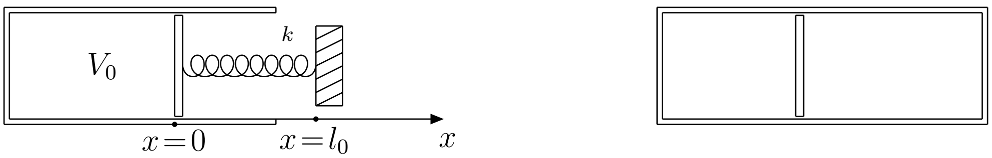

Figure 13: Left: A gas in the adiathermally isolated cylinder closed with the movable piston on the spring. Right: Two parts of the cylinder (isolated as a whole) containing two gases are separated by the adiathermal movable piston.

enclosed in two adjacent parts of a cylinder (isolated as a whole) separated by a movable but adiathermal piston[73] (Figure 13, right). As in the examples considered above, one can imagine that in equilibrium the piston stays in some fixed position and the two gases occupy the volumes $V_1^{\rm eq}$ and $V_2^{\rm eq}$ ($V_1^{\rm eq} + V_2^{\rm eq} = V$) having some well defined internal energies $U_1^{\rm eq}$ and $U_2^{\rm eq}$ (with $U_1^{\rm eq} + U_2^{\rm eq} = U$). As in the preceding example one considers then virtual equilibrium states of the system which can be obtained by blocking the (adiathermal) piston in different positions. One has also to admit different energies U_1 and U_2 (of course, respecting $U_1 + U_2 = U$) of these virtual equilibrium states because, although the adiathermal piston inhibits heat transfer between the two gases, they can do work on each other through the pressure they exert on one another through the movable piston. Thus as previously one can write as previous the condition (99)

$$\delta S = \left(\frac{1}{T_1^{\rm eq}} - \frac{1}{T_2^{\rm eq}}\right)\delta U + \left(\frac{p_1^{\rm eq}}{T_1^{\rm eq}} - \frac{p_2^{\rm eq}}{T_2^{\rm eq}}\right)\delta V\,. \tag{100}$$

However in this case here the variations δV and δU cannot be treated as independent. The point is that we now consider only those virtual equilibrium states which can be realized by blocking the piston in the real system in which any transfer of energy between the two parts must be due to the mechanical work done by the piston in the adiathermal conditions. In these conditions $\delta U = -p_2^{\rm eq}\delta V$ (the work done on the gas number 1 equals the change δV of its volume times the external pressure which is provided by the pressure of the gas number 2). Using this relation converts the condition (100) into

$$\delta S = \frac{p_1^{\rm eq} - p_2^{\rm eq}}{T_1^{\rm eq}}\,\delta V = 0\,.$$

This shows that the equilibrium position of the piston must be such as to make the pressures of both gases equal (which is anyway obvious on purely mechanical grounds), but does not impose any condition on the equilibrium temperatures $T_1^{\rm eq}$ and $T_2^{\rm eq}$. Therefore the equilibrium state of the system is not determined by the maximum entropy principle.

If a real system of this sort is prepared in a state in which the pressures of the two gases are equal it superficially seems that the temperatures T_1 and T_2 could indeed be

---

[73] A strange feature of a similar system has been demonstrated in classes.

arbitrary, but this neglects the role played by fluctuations.[74] If the initial pressures of the two gases were not equal, the piston would oscillate eternally, if there were no frictional and viscous forces. In real systems frictional and viscous forces are always present and will ultimately damp the oscillations of the piston. This first process, which lies entirely outside the domain of thermodynamics, will lead to equalization of the pressures. At the second stage, it will be the fluctuations which will lead to the equalization of the temperatures: if the temperatures of the gases are unequal, the fluctuations in two parts of the cylinder will be different and the piston itself will act as a Brownian particle receiving unequal kicks from both sides; this will result in energy transfer leading eventually to equal temperatures.

Finally we discuss the chemical equilibrium, that is, equilibrium with respect to possible matter transfer. Let n moles of a gas fill two parts (of volumes V_1 and V_2) of a cylinder separated by an unmovable wall permeable to the gas molecules. Since microscopically the heat transfer occurs through collisions of particles, such a wall is necessarily diathermal. We now seek the equilibrium distribution of the numbers $n_1^{\rm eq}$ and $n_2^{\rm eq}$ of moles of the gas and of the energies $U_1^{\rm eq}$ and $U_2^{\rm eq}$ between the two parts of the cylinder applying the maximum entropy principle. To this end we consider virtual equilibrium states realized by an adiathermal and nonpermeable wall (replacing the original one) and departures δn of the number of moles of the gas contained in the volume V_1 and δU of its energy from the equilibrium values. Since the equilibrium state maximizes entropy, the first order variation of the total entropy of the system in equilibrium

$$\begin{aligned}\delta S = {} & \left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,n_1}\Bigg|_{U_1=U_1^{\rm eq},n_1=n_1^{\rm eq}} \delta U + \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,n_2}\Bigg|_{U_2=U_2^{\rm eq},n_2=n_2^{\rm eq}} (-\delta U) \\ & + \left(\frac{\partial S_1}{\partial n_1}\right)_{U_1,V_1}\Bigg|_{U_1=U_1^{\rm eq},n_1=n_1^{\rm eq}} \delta n + \left(\frac{\partial S_2}{\partial n_2}\right)_{U_2,V_2}\Bigg|_{U_2=U_2^{\rm eq},n_2=n_2^{\rm eq}} (-\delta n)\,, \qquad (101)\end{aligned}$$

must vanish. As the departures δU and δn are independent, this entails the equality of the temperatures of the gases in the two subvolumes and the equality of their chemical potentials: $\mu_1^{\rm eq} = \mu_2^{\rm eq}$. Thus the chemical potential of a given material component (molecules of a given sort) plays the role of the indicator whether two systems will be in the chemical equilibrium with respect to the flow of this material component, when they are separated by a wall permeable to this component[75] much in the same way as the temperature T plays the role of the indicator of the possible thermal equilibrium. It follows also that the

---

[74] I quote here the explanation given by K. Rejmer in his book *Ciepło⟶Zimno*, Sorus Poznań 2013, cz"eść pierwsza: Zasady. It may be questioned - the problem of the "adiabatic piston" is a subject of an eternal discussion among thermodynamics experts and each of them has his own (very strong) opinion on its solution - but I feel it is better to outline at least one possible way of analyzing the problem than to leave it with the comment "this is a hard problem".

[75] Of course if a solid, made, say, of iron is in contact with the air, the chemical potential $\mu_{\rm Fe}$ of the air (treated as one subsystem) is (nearly) zero, similarly as is (nearly) zero the chemical potential $\mu_{\rm Air}$ of the solid. Yet no transfer of matter occurs because the iron surface should be treated as a wall impermeable to both the air and the iron.

matter transfer (of a particular matter component) can occur reversibly only between the systems having the same value of the chemical potentials (of this component) again much the same way as heat can be transferred reversibly only between bodies of (nearly) equal temperatures. Finally, in the same way as in the case of heat flow, it can be shown that in maximizing entropy matter flows from the system having the higher value of the chemical potential to the system having the lower value. If the matter transfer is reversible, the difference od the chemical potentials causing it must be infinitesimal and an infinitesimal change of either (or both) potentials is sufficient to reverse the direction of the flow of matter.

Since by assumption entropy of a system is a monotonic function of the internal energy U, the fundamental relation in the entropy representation (97) can be inverted to give

$$U = U(S, X_1, \dots, X_{o-1}, n_1, \dots, n_r)\,. \tag{102}$$

This is called the fundamental relation in the internal energy representation. If known as a function of the system's global (extensive, in the case of extensive systems) parameters it too, similarly to the relation (97), contains the complete information on the system's thermodynamic properties. Owing to the standard mathematical relation between partial derivatives of a function and of its inverse, the coefficients of the total differential (if the system is a multicomponent fluid) of U

$$dU = TdS - p\,dV + \sum_{i=1}^{r} \mu_i\,dn_i\,, \tag{103}$$

are, as we now know, the temperature, (minus) pressure and the chemical potentials. It then follows that in reversible changes of the system, in which $-p\,dV$ can be treated as a differential form $đW$ of the work (in the case of (103) the mechanical volume work) performed on the system and the terms $\mu_i\,dn_i$ can be interpreted as differential forms of works related to the changes of the amount of the i-th component in the system, the term TdS represent the heat taken by the system. Therefore integrability of the heat form $đQ$ is in this approach to 2TMDL obtained almost for free!

Let us now demonstrate the fundamental role of the fundamental relation (97) which makes it really fundamental (the fundamental form of this sentence is a joke, of course). Suppose the fundamental relation of a hypothetical simple, one-component system is God-given (or given by the statistical mechanics) in the form

$$S(U, V, n) = 3a\,(U\,V\,n)^{1/3}\,. \tag{104}$$

The system is clearly extensive, for $S(\lambda U,\,\lambda V,\,\lambda n) = \lambda\,S(U, V, n)$. Computing explicitly the derivatives (98) one finds the relations

$$\frac{1}{T} = a\,\frac{(V\,n)^{1/3}}{U^{2/3}}\,, \qquad \frac{p}{T} = a\,\frac{(U\,n)^{1/3}}{V^{2/3}}\,, \qquad -\frac{\mu}{T} = a\,\frac{(U\,V)^{1/3}}{n^{2/3}}\,.$$

Upon elimination of U, the first two of these relations yield the equation of state

$$p^2V = n\,a^3T^3\,,$$

and, moreover, the first relation alone directly yields the internal energy as a function of temperature, volume and the number of moles:

$$U(T,V,n) = a^{3/2}T^{3/2}n^{1/2}V^{1/2} = n\,a^{3/2}T^{3/2}\left(\frac{V}{n}\right)^{1/2}.$$

Entropy can be now easily obtained either as a function of T, V and n, or as a function of T, p and n:

$$S(T,V,n) = 3a^{3/2}T^{1/2}n^{1/2}V^{1/2} = 3n\,a^{3/2}T^{1/2}\left(\frac{V}{n}\right)^{1/2}, \qquad S(T,p,n) = 3n\,a^3\,\frac{T^2}{p}\,.$$

From these two forms of the entropy the two principal heat capacities C_V and C_p can readily be obtained

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_V = \frac{3}{2}\,n\,a^{3/2}T^{1/2}\left(\frac{V}{n}\right)^{1/2} = \frac{3}{2}\,n\,a^3\,\frac{T^2}{p}\,,$$

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_p = 6\,n\,a^3\,\frac{T^2}{p} = 6\,n\,a^{3/2}T^{1/2}\left(\frac{V}{n}\right)^{1/2}.$$

Of course, they satisfy the (hopefully) well-known relation (which follows from 2TMDL)

$$C_p = C_V + T\left(\frac{\partial p}{\partial T}\right)_V\left(\frac{\partial V}{\partial T}\right)_p = C_V - T\left[\left(\frac{\partial V}{\partial T}\right)_p\right]^2\left[\left(\frac{\partial V}{\partial p}\right)_T\right]^{-1},$$

because

$$-T\left[\left(\frac{\partial V}{\partial T}\right)_p\right]^2\left[\left(\frac{\partial V}{\partial p}\right)_T\right]^{-1} = -T\left(9\,\frac{n^2a^6T^4}{p^4}\right)\left(-\frac{p^3}{2na^3T^3}\right) = \frac{9}{2}\,n\,a^3\frac{T^2}{p}\,.$$

Finally, the last of the derivatives (98) gives the chemical potential which, upon eliminating U from it, can be written either as a function of T and V (in fact $v = V/n$) or T and p:

$$\mu(T,v) = -a^{3/2}T^{3/2}\left(\frac{V}{n}\right)^{1/2} = -\frac{a^3T^3}{p} = \mu(T,p)\,.$$

It is left for the students to check that the same results can be derived from the fundamental relation (102) in the energy representation which in this case reads

$$U(S,V,n) = \frac{S^3}{27a^3n\,V}\,.$$

It should be also clear that if entropy S were known as a function of T, V and n (or, analogously, internal energy U as a function of these variables), one could compute the heat capacity C_V, but there would be no way to find the equation of state or the heat capacity C_p. This illustrates the statement that $S = S(U, V, n)$ contains the complete thermodynamic information about the system while $S = S(T, V, n)$ does not.

The system considered in the example above (not a realistic one) was clearly extensive (all quantities like U, C_V, C_p when expressed in terms of $v = V/n$ were proportional to n). To stress that although the majority of realistic systems can be treated as extensive, there are nevertheless important nonextensive systems, we conclude this lecture with a brief discussion of the thermodynamics of black holes.

A digression. Black hole entropy (example of a nonextensive system).[76]

The first black hole solution of the Einstein's equations of General Relativity was obtained by K. Schwarzschild in 1916. Just before the World War II J.R. Oppenheimer and H. Snyder showed that a collapsing sufficiently massive star ends up as a black hole. At present many astrophysical objects have been identified by astronomers as black holes (and recently even gravitational waves emitted by merging two black holes have been registered by the LIGO Collaboration). In addition to the Schwarzschild solution which represents the simplest possible such object characterized entirely by its total mass M, there also solutions representing rotating and/or electrically charged black holes (known as the Kerr, Reisner-Nordström and Kerr-Newman solutions, respectively) which in addition to the mass are characterized by their total charge Q and/or angular momentum $\mathbf{L}$. It is known that M, Q and $\mathbf{L}$ are the only possible quantities which can characterize a black hole (famous saying that “black holes have no hairs”).

Initially a black hole was viewed as an entirely passive object which cannot emit anything - it can only absorb. If this were true, black holes would lie outside the range of applicability of thermodynamics because, as was noted by J.A. Wheeler, 2TMDL would not apply to systems including black holes: if a portion of matter of nonzero entropy were dropped into a black hole, the external observer could not be sure that the total entropy of the system had not decreased, because the only characteristics of the black hole available to him, M, Q and $\mathbf{L}$, cannot tell how much entropy the black hole has swallowed. This has led J.D. Bekenstein,[77] a Wheeler's student, to formulate in 1972 the conjecture that the black hole has in fact a nonzero entropy which is proportional to the area A of its horizon surface (a surface separating from the rest of the space the region from which light or any other object cannot escape), which in turn is a function of M, Q and $\mathbf{L}$. This was prompted by a theorem of S. Hawking which says that according to the laws of classical General Relativity, the total area A of horizons of black holes cannot decrease and it increases in dynamical processes like e.g. merging of two black holes. Another hint was the observation made by D. Christodoulou (another student of Wheeler): investigating the efficiency of the so called Penrose process by which energy related to the rotation can

---

[76] Based on the article by J.D. Bekenstein *Physics Today*, **33**, 24 (1980); another standard introductory reference is: B.R. Parker and R. McLeod *Am.J.Phys.* **48** 1066 (1980); this second article is however not very useful - its authors evidently had problems with clear formulation of their thoughts.

[77] Just a year, or so, I saw his obituary in the CERN Courier.

be extracted from a rotating black hole he found that it is the greatest if the process is realized reversibly and that in this case the area A of the black hole horizon does not increase - this indicated that some sort of thermodynamics might be at play...

The general formula proposed by Bekenstein (the coefficient 1/4 has been fixed later by Hawking) which gives the black hole entropy reads:

$$S = \frac{1}{4}\, k_{\rm B}\, \frac{A}{\ell_{\rm Pl}^2}\,. \tag{105}$$

$k_{\rm B}$ is the Boltzmann constant and $\ell_{\rm Pl} \equiv \sqrt{\hbar G/c^3}$ is the Planck length ($G = 6.67\times 10^{-11}$ ${\rm m}^3\,{\rm kg}^{-1}{\rm sec}^{-2}$ is the gravitational constant; $\ell_{\rm Pl} = 1.6\times 10^{-35} m$). It is to be noted that this formula relates a purely thermodynamical quantity S to a purely gravitational one (the horizon area) and involves the Planck constant signaling existence of a deep connection between thermodynamics, gravitation and quantum phenomena. The formula for the horizon area A of the Schwarzschild black hole ($Q = 0$, $\mathbf{L} = \mathbf{0}$) is very simple:

$$A_{\rm Sch} = 4\pi R_g^2 = 16\pi\, \frac{G^2}{c^4}\, M^2 = 16\pi\, \frac{\hbar G}{c^3}\, \frac{M^2}{M_{\rm Pl}^2} = 16\pi \ell_{\rm pl}^2\, \frac{M^2}{M_{\rm Pl}^2}\,. \tag{106}$$

here $M_{\rm Pl} = \sqrt{\hbar c/G} = 2.2\times 10^{-8}kg= 1.22\times 10^{19}$ GeV$/c^2$ is the Planck mass. $R_g = 2GM/c^2$ is just the radius which can be obtained using the Newtonian dynamics: it is the radius of a planet of mass M for which the second cosmic speed v_{II} (determined by the condition $\frac{1}{2} m v_{II}^2 - GMm/r = 0$) needed to reach infinity starting from its surface is equal c - since this is the highest possible speed, R_g is just the radius of the horizon (bodies closer to the center than R_g cannot escape). Thermodynamics of the Schwarzschild black hole is very simple:

$$\begin{aligned} S &= 4\pi k_{\rm B}\, \frac{M^2}{M_{\rm Pl}^2}\,, \\ U &= Mc^2\,. \end{aligned} \tag{107}$$

This is in fact the simplest of all thermodynamical systems since its equilibrium state is characterized by only one parameter M: no extraction of energy from it by a reversible work is possible (the number $o - 1$ of its deformative parameters is zero). The equation

$$S(M) = 4\pi k_{\rm B}\, \frac{U^2}{M_{\rm Pl}^2 c^4}\,, \tag{108}$$

plays the role of the fundamental relation (97) of the Schwarschild black hole thermodynamics. It is to be noted that in this case S is *not* a homogeneous first order function of its only global argument U. The simple scaling is violated by the presence of the fundamental mass (or length scale) $M_{\rm Pl}$ (or $\ell_{\rm Pl}$). Applying to (108) the standard formula $dS = dU/T$ one gets

$$\frac{1}{T} = \frac{dS}{dU} = \frac{8\pi k_{\rm B}}{M_{\rm Pl}^2 c^4}\, U\,, \tag{109}$$

or

$$U = \frac{M_{\rm Pl}^2 c^4}{8\pi k_{\rm B} T}\,, \tag{110}$$

The resulting formula for the Schwarzschild black hole temperature can also be written in the form

$$T = \frac{\hbar}{2\pi k_{\rm B} c}\,\frac{M_{\rm pl}^2 c^3}{4\hbar M} = \frac{\hbar}{2\pi k_{\rm B} c}\left(\frac{c^4}{4GM}\right) \equiv \frac{\hbar}{2\pi k_{\rm B} c}\,\kappa\,, \tag{111}$$

in terms of the so-called surface gravity κ which is the "g" (that is $g = GM/R_Z^2$ known from ordinary mechanics at the Earth surface) of a planet of mass M and radius R_g. It follows that the greater is the mass (the internal energy U) of the black hole, the lower is its temperature! This of course results in the negative its heat capacity

$$C = T\,\frac{dS(U(T))}{dT} = \frac{dU}{dT} = -\frac{k_{\rm B}}{8\pi}\left(\frac{M_{\rm Pl} c^2}{k_{\rm B} T}\right)^2. \tag{112}$$

If a black hole has nonzero angular momentum $\mathbf{L}$ and/or electric charge Q, extraction of energy from it by a reversible mechanical or electrical process is possible (we have already mentioned the mechanical Penrose process) by coupling to $\mathbf{L}$ or to Q appropriate external agents. Such reversible works can be continued until $\mathbf{L} = \mathbf{0}$, $Q = 0$. The formula for the area of the Kerr-Newman black hole horizon surface is best expressed in terms of the parameters (which both have dimension of mass squared; we use the normal Gauss system of units so that $Q^2/\hbar c$ is dimensionless - $e^2/\hbar c = \alpha_{\rm EM}$ as every high energy physicist knows)

$$\begin{aligned} \tilde{Q}^2 &\equiv \frac{Q^2}{G} = \frac{Q^2}{\hbar c}\,\frac{\hbar c}{G} = \frac{Q^2}{\hbar c}\,M_{\rm Pl}^2\,, \\ \mathbf{a}^2 &= \frac{c^2}{G^2}\,\frac{\mathbf{L}^2}{M^2} = \frac{\mathbf{L}^2}{\hbar^2}\left(\frac{\hbar c}{G}\right)^2 \frac{1}{M^2} = \frac{\mathbf{L}^2}{\hbar^2}\,\frac{M_{\rm Pl}^4}{M^2}\,, \end{aligned}$$

and reads

$$A = 4\pi\,\frac{G^2}{c^4}\left[\left(M + \sqrt{M^2 - \tilde{Q}^2 - \mathbf{a}^2}\right)^2 + \mathbf{a}^2\right]. \tag{113}$$

This formula together with (105) and the identification $U = Mc^2$, gives the fundamental Callenian relation of the most general black hole thermodynamics:

$$S = S(U, \mathbf{L}, Q)\,. \tag{114}$$

It is again clear that the black hole is not an extensive system. All thermodynamical information can be obtained from it in the standard way. In particular

$$dS = \frac{1}{T}\,dU - \frac{\phi}{T}\,dQ - \frac{1}{T}\,\mathbf{\Omega}\cdot d\mathbf{L}\,, \tag{115}$$

which leads to the form

$$dU = TdS + \phi\, dQ + \mathbf{\Omega}\cdot d\mathbf{L}\,, \tag{116}$$

of 1TMDL as applied to all differential changes of the black hole.[78] ϕ is here the electrostatic potential and $\mathbf{\Omega}$ the black hole's angular velocity. The second and the third term on the right hand side represent infinitesimal reversible works that can be done on or by the black hole. The rest is a matter of ordinary thermodynamical fiku-miku, which although can be somewhat complicated because of the complicated form of the relation (113), should not be more complicated from the point of view of principles, than any other thermodynamical problem. A curious student may try to compute for instance the heat capacity C_ϕ of the Reisner-Nordström ($\mathbf{L} = \mathbf{0}$) black hole or C_Ω of the Kerr ($Q = 0$) one.

If the black hole parameter T is to be regarded as the true temperature (and not merely as an analog of it), the black hole should, as every body at a nonzero temperature T, radiate.[79] Classically a black hole cannot radiate, because nothing, including light, can escape from it. But the presence of the Planck constant $\hbar$ in the Bekenstein formula (105) strongly suggests, as has already been stressed, that the connection between gravity and thermodynamics is not based on classical physics. It was Hawking who by considering quantum processes occurring near the horizon of a forming Schwarzschild black hole[80] discovered (in 1974) that it radiates just as does a body at temperature T given by (111). In this process the black hole horizon decreases (which classically is impossible), the black hole "evaporates" and its entropy decreases. This decrease of entropy is however overcompensated (if the evaporation occurs in the surrounding of lower temperature) by the entropy of the emitted thermal radiation (which consist of all kinds of elementary particles). In this way the entropy law and thermodynamics as such are applicable also to black holes.

---

[78] But of course, it is only in reversible changes that ϕdQ and $\mathbf{\Omega}{\cdot}d\mathbf{L}$ have respectively the interpretations of the works done on the black hole electrically and mechanically.

[79] This is a point usually left aside in discussing equilibrium states of thermodynamical systems: to be truly isolated adiathermally, a system must be placed in a shield which not only isolates it from all influences from without including any incoming radiation, but must also ideally reflect back the thermal radiation emitted by the system itself, when its temperature is not zero.

[80] It is these considerations which allowed Hawking to fix the coefficient in the Bekenstein formula (105).

LECTURE VII (TMD)

We now proceed to exploiting the entropy law and other properties of this quantity. We will first consider the problem of maximal useful work which can be extracted from a given system in given conditions (i.e. from a system with given constraints). These considerations will lead us to the formulation of the general principle which determines the equilibrium states of systems open to (i.e. in contact with) their surroundings, that is systems in conditions in which directly controlled are not necessarily (all) their deformative parameters $X_1, \dots, X_{o-1}$ and energy U but rather some of the generalized forces $y_1, \dots, y_{o-1}$ and/or the temperature T. This principle will generalize the Callenian maximum entropy principle which applies to systems which are not in contact with their surroundings. They will also allow to formulate the general conditions of stability of thermodynamic systems.

We begin by showing that entropy of an extensive system is a concave (concave upwards) function of any of its extensive arguments. Indeed, consider two systems A and B both isolated and in each equilibrium within itself. Let us write their entropies as

$$S_A = S_A(\lambda_A U_A,\, \xi_A V_A, \dots)\,, \qquad S_B = S_B(\lambda_B U_B,\, \xi_B V_B, \dots)\,, \tag{117}$$

that is, the actual values of the extensive parameters of the two systems have been written as other values of these parameters scaled up by the respective arbitrary factors λ, ξ, etc. We now imagine that the two systems (treated as a compound isolated system) have been brought into thermal contact and a new equilibrium state of the compound system has been reached. By the entropy law, the entropy of this new equilibrium state, $S_{A+B}(\lambda_A U_A + \lambda_B U_B,\, \xi_A V_A + \xi_B V_B, \dots)$ cannot be smaller than the sum of the entropies S_A and S_B:

$$S_{A+B}(\lambda_A U_A + \lambda_B U_B,\, \xi_A V_A + \xi_B V_B, \dots) \geq S_A(\lambda_A U_A, \xi_A V_A, \dots) + S_B(\lambda_B U_B, \xi_B V_B, \dots)\,,$$

irrespectively of the (positive) values of the factors λ, ξ, etc. and irrespectively of the values of their energies, volumes, etc. If the two systems are of the same kind (e.g. two pieces of the same solid), the functions S_A and S_B are given by the same function $S(\cdot, \cdot, \dots)$ but evaluated at different values of its arguments. If in addition one sets $\lambda_A = \xi_A = \lambda$, $\lambda_B = \xi_B = 1 - \lambda$, etc. and considers $V_A = V_B = V$, etc., one obtains the inequality

$$S(\lambda U_A + (1-\lambda) U_B\,, V, \dots) \geq \lambda\, S(U_A, V, \dots) + (1-\lambda)\, S(U_B, V, \dots)\,,$$

which just means concavity upwards of the entropy $S(\cdot, \cdot, \dots)$ of the extensive system as a function of its first (energy) argument (Figure 14). In the same way one can demonstrate its concavity as a function of its volume argument and all other extensive deformative parameters.

Concavity upwards of the entropy as a function of the internal energy is necessary for stability of the system: indeed, if the entropy were not concave upwards it could assume greater values in an inhomogeneous state than in the homogeneous one, as an example

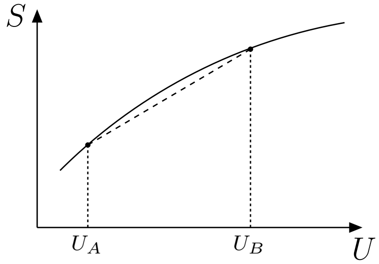

Figure 14: Entropy of an extensive system is a concave upwards function of its internal energy. The dashed line shows the function $\lambda S(U_A) + (1-\lambda)S(U_B)$ as a function of the variable $\lambda U_A+(1-\lambda)U_B$. Everywhere between U_A and U_B the value of $S(\lambda U_A+(1-\lambda)U_B)$ is above this line.

presented in Callen's textbook shows. Global concavity implies the local one which means that

$$\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,\dots} < 0\,. \tag{118}$$

Because of the first of the equalities (98) this amounts to

$$\left(\frac{\partial}{\partial U}\frac{1}{T}\right)_{V,\dots} = -\frac{1}{T^2}\left(\frac{\partial U}{\partial T}\right)^{-1}_{V,\dots} \equiv -\frac{1}{T^2 C_V} < 0\,, \tag{119}$$

which in turn, in view of the positivity of the factor T^2, means that $C_V > 0$. We have seen that the black hole analog of C_V is negative because the black hole is not an extensive system.

S and U as thermodynamical potentials
We will now show that the postulate that the equilibrium state of an isolated system maximizes the system's entropy on the set of all virtual equilibrium states (which can be fabricated with the help of constraints stronger than the ones the system is actually subjected to) at a fixed value of its internal energy U (and fixed values of other global parameters characterizing the system as a whole) - **the maximum entropy principle** - is equivalent to the statement that the equilibrium state of an isolated system minimizes the internal energy on the set of all virtual states at the fixed value of the system's entropy S (and fixed values of other global parameters) - this will be called **the minimum energy principle**. This second statement sounds a little bit more abstract - it is easier to imagine keeping the system's internal energy fixed (and distributing it among its subsystems) than to imagine a fixed value of the system's entropy and varying the internal energy. Nevertheless, the two formulations are equivalent. We will as usually give two proofs: one physical and general (but requiring some imagination) and another one mathematical, more concrete, but necessarily restricted to a specific case.

To prove that the maximum entropy principle implies the minimum energy principle, we make the *reductio ad absurdum* and assume that the equilibrium state (corresponding to given constraints) in which the entropy of the system is S_0 and its energy is U_0 does not minimize energy; there should therefore exist another equilibrium state (which can be blocked with the help of stronger constraints) in which the system's entropy is still equal S_0 but its energy U is lower than U_0. Therefore it would be possible to take the system adiathermally and reversibly (adiabatically) to this state of lower energy (because the two states have the same entropy S_0, it logically should be possible to achieve this by e.g. suitable Carnot cycles run between different parts of the system playing the roles of heat reservoirs) extracting from it the work $\bar{W} = U_0 - U$. The energy gained in this way in the form of a mechanical work could be then put back irreversibly into the system (temporarily breaking its isolation from the surrounding and then restoring it back) bringing it into an equilibrium state of energy U_0 but of entropy higher than S_0. In effect, the equilibrium state would not maximize entropy at fixed value of the energy because there would be a state (possibly with stronger constraints) of the same energy and higher entropy.

The proof of the converse assertion (that the minimum energy principle implies the maximum entropy principle) goes similarly: we assume that the system's equilibrium state (corresponding to given constraints) of entropy S_0 and energy U_0 does not maximize entropy (on the set of virtual states with stronger constraints but the same energy U_0), so there should exist a state of higher entropy $S > S_0$ and the same energy U_0 (realized when the system is subjected to stronger constraints); the system in this higher entropy state could be then brought into thermal contact (breaking its isolation and then restoring it back) with a reservoir (a heat bath) transferring to it some heat; entropy of the system would be thereby lowered to S_0 (things could be arranged so) and, by 1TMDL, also the energy of the system would be lowered. One would therefore end up with the system in equilibrium with the entropy S_0 and energy lower than U_0 (and stronger constraints). The assumed equilibrium state of entropy S_0 and energy U_0 would not, therefore, be the state of lowest possible energy (at fixed value S_0 of entropy).

The mathematical reasoning requires specifying a parameter, let's call it x, (or a couple of parameters) which represents (represent) changes of the constraints imposed on the system; one can assume that $x = 0$ corresponds to the actual constraints to which the system is subjected and $x \neq 0$ corresponds to stronger constraints. The entropy maximum principle means then that

$$\left(\frac{\partial S}{\partial x}\right)_{U,\dots}\bigg|_{x=0} = 0\,, \qquad \left(\frac{\partial^2 S}{\partial x^2}\right)_{U,\dots}\bigg|_{x=0} < 0\,. \tag{120}$$

Using the "shocking" relation one can write

$$\left(\frac{\partial U}{\partial x}\right)_{S,\dots}\bigg|_{x=0} = -\left(\frac{\partial U}{\partial S}\right)_{x,\dots}\left(\frac{\partial S}{\partial x}\right)_{U,\dots}\bigg|_{x=0} = 0\,. \tag{121}$$

This shows that (except at $T = 0$) the derivative $(\partial U/\partial x)_S$ vanishes in the same state (corresponding to $x = 0$) in which $(\partial S/\partial x)_U$ vanishes. The second derivative of U consists of two terms

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_{S,\dots} = -\left(\frac{\partial S}{\partial x}\right)_{U,\dots}\left(\frac{\partial^2 U}{\partial S\,\partial x}\right) - \left(\frac{\partial U}{\partial S}\right)_{x,\dots}\left[\frac{\partial}{\partial x}\left(\frac{\partial S}{\partial x}\right)_{U,\dots}\right]_{S,\dots}.$$

The first term vanishes at $x = 0$; in the second term the derivative $D \equiv (\partial S/\partial x)_{U,\dots}$ which is a function of the variables x and U should be now treated as $D(U(x, S),\, x)$ so that

$$\left(\frac{\partial D}{\partial x}\right)_{S,\dots} = \left(\frac{\partial D}{\partial x}\right)_{U,\dots} + \left(\frac{\partial D}{\partial U}\right)_{x,\dots}\left(\frac{\partial U}{\partial x}\right)_{S,\dots}.$$

Again, the second term vanishes at $x = 0$ (because $(\partial U/\partial x)_{S,\dots} = 0$ there) and one ends up with

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_{S,\dots} = -T\left(\frac{\partial^2 S}{\partial x^2}\right)_{U,\dots},$$

which shows that if $S(U, x, \dots)$ has a maximum at $x = 0$, then $U(S, x, \dots)$ has there a minimum. With a little bit mor labour the proof can be extended to more parameters representing constraints, but its weakness is that one considers the same constraints for fabricating virtual states at constant energy and at constant entropy.

Either way, the conclusion is that with respect to isolated systems entropy and internal energy play the roles of thermodynamical potentials which (in the sense discussed in this and in the preceding Lectures) determine the system's equilibrium states.

Maximal and Minimal work

One of the main applications of the entropy law is putting an upper limit on useful work which can be extracted from a given thermodynamical system by changing appropriately (weakening) the constrains to which the system is subjected, thereby allowing it to reach another equilibrium state. The converse problem is putting a lower limit on the work necessary to bring a given system from one equilibrium state to another one (with stronger constraints).

The simplest situations in which the problem of the maximal (minimal) work can be posed and analyzed is when a given system consisting of several subsystems can only exchange mechanical work with an external "mechanical work source". In particular, it is as a whole adiathermally isolated and the total amount of matter in it is not changing. The internal constraints of the system (which keep its different parts in equilibrium within themselves) can be weakened and on the way to the equilibrium state corresponding to the new constraints the system can yield some work; since the system is not completely isolated[81] (it can exchange work with the surrounding), the final equilibrium state and the work extracted from the system depend on the process by which the new equilibrium state

---

[81] Recall that if the system were completely isolated, its final state would be uniquely determined by the new constrains. In all reasonings which follow it is important to keep in mind the distinction between a concrete (equilibrium) final state (compatible with the given final constraints) and a set of possible final states specified only by prescribing the final constraints.

is reached. The questions is, which process will result in the maximal extracted work. In analyzing this we will assume that the final and initial total volume of all subsystems comprising the system is the same.[82] Of course, it can vary during the process. In addition, some subsystems of the considered system may play only a subsidiary roles in the sense that their initial and final states are the same.

Since the system is supposed not to exchange heat with the surrounding, by 1TMDL, the work $\bar{W}$ extracted from it is just

$$\bar{W} = U_{\rm in} - U_{\rm fin}\,. \tag{122}$$

The energy $U_{\rm in}$ is fixed but the final state and, therefore, $U_{\rm fin}$ depends on the process. It is, however, a function of the final state entropy. Since from the Callenian postulates U is a growing function of entropy (recall the assumption $(\partial S/\partial U)_X > 0$ which implies $(\partial U/\partial S)_X > 0$), the greatest maximal work is obtained if entropy of the system does not increase that is, if the process by which the system reaches the final equilibrium state corresponding to the new weakened constraints is reversible.

The standard illustration of this sort of situations is the system consisting of two bodies of unequal temperatures, say T_1 and $T_2 > T_1$, separated by an adiathermal wall. The final constraint is a diathermal wall through which the two bodies are in thermal contact and have, hence, the same final temperature. If the two bodies are brought into thermal contact directly (one possible process of attaining the final equilibrium state), the entropy of the whole system will increase (we talked about this in Lectures V and VI) and no useful work at all will be extracted from the system. The final common temperature $T_{\rm fin}$ of the bodies will be determined by 1TMDL which, assuming for simplicity[83] that the heat capacities C at constant volume[84] of the two bodies are equal and (to make things simple) independent of the temperature, takes the form:

$$U_{\rm in} = (U_0 + CT_1) + (U_0 + CT_2) = (U_0 + CT_{\rm fin}) + (U_0 + CT_{\rm fin}) = U_{\rm fin}\,, \tag{123}$$

that is, $T_{\rm fin} = \frac{1}{2}(T_1 + T_2)$. The entropy change in such a process will be

$$\Delta S = 2C\ln\frac{T_{\rm fin}}{T_0} - C\ln\frac{T_1}{T_0} - C\ln\frac{T_2}{T_0} = C\ln\frac{T_{\rm fin}^2}{T_1T_2} = C\ln\frac{(T_1+T_2)^2}{4T_1T_2} > 0\,. \tag{124}$$

If, however, the process by which the two bodies attain thermal equilibrium with one another is reversible, entropy will not change and it is the condition $\Delta S = C\ln(T_{\rm fin}^2/T_1T_2) =$

---

[82] In this way we exclude the trivial work the system could do changing its volume - we are interested only in the work which can be obtained by changing the internal constraints of the system: allowing it to change its total volume would mean changing its external constraints. In fact, the condition of equal initial and final total volumes can be weakened by assuming that the pressure p_0 of the surrounding is negligible (or zero) so that changing its total volume the system does not do any work. This will be taken into account when the general case is considered.

[83] The most general case is treated in one of the homework Problems.

[84] As said, one could also consider the heat capacities at constant internal pressure if the external pressure is zero.

0 which will determine the final temperature $T_{\rm fin} = \sqrt{T_1 T_2}$. By 1TMDL the work extracted from the system in this case will be

$$\begin{aligned}\bar{W} &= U_{\rm in} - U_{\rm fin} = 2U_0 + C(T_1 + T_2) - 2\left(U_0 + C\sqrt{T_1 T_2}\right) \\ &= C\left(T_1 + T_2 - 2\sqrt{T_1 T_2}\right) = C\left(\sqrt{T_2} - \sqrt{T_1}\right)^2 . \end{aligned} \tag{125}$$

One can of course ask how to realize "practically" such a reversible process? In this case (but not necessarily always) the answer is simple: it suffices to run a Carnot engine treating the body at T_2 as the hotter (higher temperature) heat reservoir and the body at T_1 as the colder one. In the calculation one has of course to take into account that the exchange of heat between the bodies and the engine lowers the temperature of the hotter reservoir and raises the one of the colder one, so that finally they will equalize. If the actual (during the working of the engine) temperature of the hotter reservoir is T_+ ($T_{\rm fin} \leq T_+ \leq T_2$) and that of the colder one is T_- ($T_1 \leq T_- \leq T_{\rm fin}$) then (upon using the actual efficiency $\eta(T_-, T_+) = 1 - T_-/T_+$ of the Carnot cycle working between the temperatures T_- and T_+)

$$đ\bar{W} = \eta(T_-, T_+)\, đQ_2 = \left(1 - \frac{T_-}{T_+}\right)(-C\, dT_+)\,,$$

and the necessary correlation of T_- with T_+ is provided by the condition

$$\frac{đQ_2}{T_+} + \frac{đQ_1}{T_-} = -C\left(\frac{dT_+}{T_+} + \frac{dT_-}{T_-}\right) = 0\,,$$

of constancy of the entropy ($đQ_2 = -C dT_+ > 0$ and $đQ_1 = -C dT_- < 0$ are the heats taken by the engine from the two bodies) which upon integration with the obvious initial conditions yields $T_- = T_1 T_2 / T_+$ (and, of course, $T^2_{\rm fin} = T_1 T_2$). Integrating then $đW$ from T_2 to $T_{\rm fin}$ gives the same result (125).

The converse situation to the one analyzed above is when the system is supposed to attain, as a result of the change, another equilibrium state of higher energy corresponding to stronger constraints. The analogous reasoning then shows that the minimal work W which must be done on the system corresponds to an isentropic change.

A more general situations is when a system, which can be composed of several subsystems, can exchange with the surrounding not only work (with the mechanical work source, MWS) but also heat and perhaps also matter; one can also admit that the total volume of (all parts of) the system can be different in the final state than in the initial state. To analyze such situations in general terms one can model the system's surrounding as consisting of a reversible[85] heat source (RHS) - a reservoir so large that it always remains in equilibrium at temperature T_0, no matter how big finite amount of heat is

---

[85] "Reversible" in the names of the source is to mean that the they remain in equilibrium in themselves; it is not meant to imply that the exchanges batween them and the system are reversible in the usual sense.

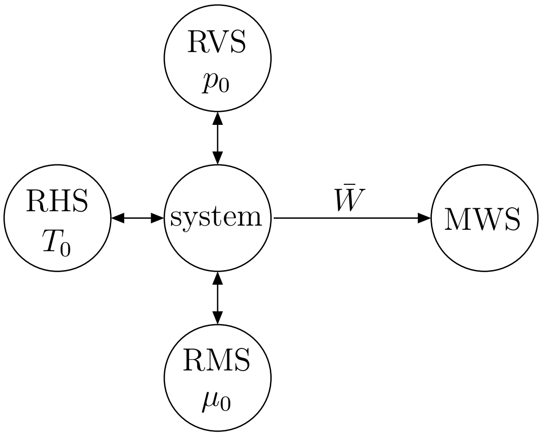

Figure 15: The system and the set of reservoirs representing its surrounding. The parameters T_0, p_0, μ_0 of the sources modeling the system's surrounding are constant.

extracted from or added to it, a reversible volume source (RVS) - a very large system in equilibrium (the entropy of which stays always constant and need not be taken into account) at invariable pressure p_0 and, finally, a source of matter[86] (RMS) in equilibrium (the entropy of which also stays always constant and need not be taken into account - recall that the chemical potential is defined by (98) as the change of energy of the body, here the RMS, resulting from changing the amount of matter in it at constant entropy) at invariable chemical potential μ_0 (this general setting is schematically illustrated in Figure 15). Owing to the constraints - the internal ones and also the external ones, separating it from its surrounding - the system, which may consist of several subsystems, is initially in equilibrium (its individual subsystems can have arbitrary different temperatures, pressures, chemical potentials). If the constraints are changed (are weakened or strengthened), the final equilibrium state compatible with the new constraints can depend on the process by which the system attains it. In the process of reaching this new equilibrium state the system may (if the external constraints are weakened appropriately) exchange heat with the RHS. We also assume that any change of the total volume of the system (the sum of volumes of its subsystems) is compensated by the opposite change of the volume of the RVS which is due to a mechanical contact between the parts of the system and the RVS. Similarly, any change in the total matter content of the system is compensated by the opposite change of matter content of the RMS due to a direct flow of matter between parts of the system and the RMS.[87]

---

[86] If there is more than one matter component which can be exchanged with the surrounding, one can introduce one such source per component.

[87] It may seem that if there were more matter constituents in the system (and correspondingly more matter sources representing the surrounding) and chemical reactions were allowed to occur in the system

By 1TMDL, the work extracted (by MWS) from the entire supersystem: the system itself and its surrounding represented by RHS, RVS and RMS, is

$$\bar{W} = -(\Delta U + \Delta U_{\rm RHS} + \Delta U_{\rm RVS} + \Delta U_{\rm RMS})\,, \tag{126}$$

where $-\Delta U = U^{\rm in} - U^{\rm fin}$ etc. But since the reservoirs stay in equilibrium, $\Delta U_{\rm RHS} = T_0 \Delta S_{\rm RHS}$, $\Delta U_{\rm RVS} = -p_0 \Delta V_{\rm RVS}$ and $\Delta U_{\rm RMS} = \mu_0 \Delta n_{\rm RMS}$. Furthermore, by assumption $\Delta V_{\rm RVS} = -\Delta V$ and $\Delta n_{\rm RMS} = -\Delta n$ (the total volume and the total amount of matter in the system and its surrounding remain constant)

$$\bar{W} = -\left(\Delta U + T_0 \Delta S_{\rm RHS} + p_0 \Delta V - \mu_0 \Delta n\right).$$

As far as the exchange of heat (which need not be reversible) between the system and the RHS is concerned, it follows from the entropy law (applied to the isolated supersystem consisting of the system itself and the sources) that $\Delta S_{\rm RHS} + \Delta S \geq 0$, or that

$$-T_0 \Delta S_{\rm RHS} \leq T_0 \Delta S\,. \tag{127}$$

Therefore the work $\bar{W}$ which can be extracted from the system accomplishing a concrete change (specified by the total changes ΔU, ΔV, Δn of all its parts) in the given surrounding (characterized by T_0, p_0 and μ_0) is limited from above:[88]

$$\bar{W} \leq -\Delta\left(U - T_0 S + p_0 V - \mu_0 n\right) = -\Delta A\,. \tag{128}$$

The quantity $A \equiv U - T_0 S + p_0 V - \mu_0 n$ is called the **availability of the system**. Contrary to this name, it is the property of the system *and* its surrounding. Recall also once more

---

(as a result of weakening its internal constraint) and to change its matter composition, one would have to distinguish the changes of the number of moles of each of the constituents occurring due to chemical reactions ($\Delta n_i^{\rm ch}$) and occurring due to the exchanges with the matter sources in the surrounding ($\Delta n_i^{\rm inflow}$). This is not so, for reactions inside the system (as will be seen in the following Lectures) occur so that the differential (in which $dn_i^{\rm ch}$ are the resulting from reactions changes of the numbers of moles of different chemical constituents of the system between two infinitesimally close equilibrium states)

$$\sum_{i=1}^{r} \mu_i \, dn_i^{\rm ch}\,,$$

vanishes. This is consistent with the possibility of treating the system as a "black box" and controlling (by controling e.g. the total mass of the system) only the matter which enters into it. Thus Δn or dn (Δn_i or dn_i) in the formulae below represent only changes of the amount of matter due to the contact of the system with RMS. Of course, if the system does not exchange matter with the surrounding and the term $\Delta U_{\rm RMS}$ in (126) is absent one can set $\mu_0 = 0$ in all further formulae.

[88] The following point of view on the conducted reasoning may be helpful: through $\Delta V_{\rm RVS} = -\Delta V$ and $\Delta n_{\rm RMS} = -\Delta n$ one assumes that the changes of energies of the volume and matter sources (RVS and RMS) are controlled by measuring the changes (ΔV and Δn) of the system itself; in contrast, changes of the energy of the heat reservoir (RHS) are not fully controlled in this way: they can only be bounded by (127). Therefore, even if one considers a concrete final state of the system itself, the final state of the "supersystem" consisting of the system and its surrounding is not fully specified and for this reason (128) is only an inequality. Of course, still different final states of the system itself (i.e. different values of ΔU, ΔS, ΔV and Δn) can be compatible with a given set of new constraints.

that in general the quantities U, V and S pertaining in (128) to the system are the sums of energies, volumes and entropies of its different parts (subsystems).

If the change of the system from an equilibrium state to another one is infinitesimal, replacing ΔU by $dU = TdS - p\,dV + \mu\,dn$ one gets[89]

$$\bar{w} \leq -(T-T_0)dS + (p-p_0)dV - (\mu-\mu_0)dn\,. \tag{129}$$

This formula can be better understood by noting that if the change is reversible, $-TdS$ is the heat $\text{đ}Q$ lost by the system; its conversion into work absorbed by the MWS in the surrounding at temperature T_0 (e.g. by running a suitable Carnot engine) yields the useful work equal $\eta(T_0,T)\text{đ}Q = -(T-T_0)dS$. Similarly, if the system expands reversibly changing its volume by dV, the additional pressure $p-p_0$ must be applied externally (that is, $p-p_0$ is the lacking pressure necessary to make the expansion reversible in order to fully profit from it - recall the discussion in Lecture II) by external forces provided by the MWS and the work done on the MSW by the system in its expansion is just $(p-p_0)dV$. Finally, to make the change of matter content of the system by $-dn$ reversibly requires the additional chemical potential $\mu-\mu_0$ be provided by the MWS[90] and the work done by the system on MWS due to the matter flow is $-(\mu-\mu_0)dn$.

Conversely, the minimal work W needed to accomplish a given change of the system (strengthening the constraints it is subjected to) characterized by the total changes ΔU, ΔV and Δn of all its parts) in the given surrounding is bounded from below

$$W \geq \Delta\left(U - T_0S + p_0V - \mu_0 n\right) = \Delta A\,. \tag{130}$$

The minimal work $W_{\text{min}} = \Delta A$ needed to accomplish the change is given an interpretation in Figure 16 which sketches the dependence of the total entropy of the supersystem consisting of the system itself and of its surrounding (RHS, RVS and RMS) on its total energy with weaker (solid line) and stronger (dashed line) constraints imposed on the system; the horizontal dotted line shows the change of the total energy needed to strengthen the constraints at fixed entropy (it is thus equal W_{min}); the vertical dotted line shows the

---

[89] We restrict ourselves here to a system consisting of a single homogeneous body or to a situation in which all parts of the system have the same temperature T pressure p and chemical potential μ; in the general case of a system consisting of several homogeneous bodies each of which could have (owing to internal constraints) different temperatures, pressures and chemical potentials the formula (129) would take the form

$$\bar{w} \leq -\sum_a [(T_a - T_0)dS_a + (p_a - p_0)dV_a - (\mu_a - \mu_0)dn_a]\,,$$

where the sum is taken over different parts of the system.

[90] Here we extend a bit the meaning of the "work source", allowing it to change its energy also in the form of a matter flow. If one keeps the original definition of MWS, a reversible change of matter content of the system requires its chemical potential μ to be equal to μ_0 and the last term in (129) vanishes; if the matter flow between the system and RMS is not reversible, TdS does not represent the heat transfer to/from the system and $(p-p_0)dV$ does not represent the work done by it (recall the discussion at the end of Lecture IV) and (129) is valid only as the inequality.

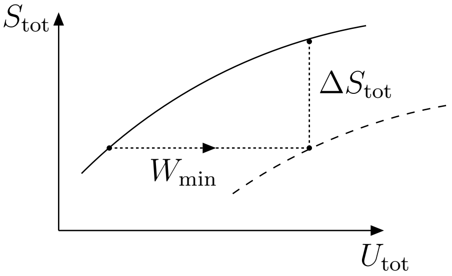

Figure 16: Interpretation of the minimal work in terms of the entropy. The solid line represents the total entropy of the system and its surrounding as a function of the total energy with the initial (weaker) constraints while the long-dashed line represents the total entropy as a function of the total energy with the final (stronger) constraints.

increase $\Delta S_{\mathrm{tot}} > 0$ of the supersystem's total entropy if the stronger constraints are replaced by the weaker ones and the system and its surrounding attain equilibrium without any intervention from without (i.e. from MWS). If the system is very small compared to the surrounding (treated here as a real system and not as an infinitely large one), then $W_{\mathrm{min}} \ll U_{\mathrm{tot}}$ and the increase of the total entropy (occurring in the spontaneous reaching by the system the equilibrium after weakening the constraints) can be written as (here U_{tot} and S_{tot} stand for the total energy and entropy of the system and of the surrounding with the weaker constraints)

$$\Delta S_{\mathrm{tot}} \approx \left(\frac{\partial S_{\mathrm{tot}}}{\partial U_{\mathrm{tot}}}\right)_{V_{\mathrm{tot}},n_{\mathrm{tot}}} W_{\mathrm{min}} = \frac{1}{T_0}\left(\Delta U - T_0\Delta S + p_0\Delta V - \mu_0\Delta n\right). \tag{131}$$

This formula quantifies by how much the total entropy S_{tot} of the system and its surrounding (treated together as an isolated supersystem) with given (stronger) constraints differs from the maximal entropy obtained when the system is in equilibrium with its surrounding (without walls blocking the direct contact between them).

The formula (128) is very general and applies to most of the situations in which the system exchanges heat, work and possibly matter with the surrounding. A few special cases deserve to be discussed separately:

- The formula (128) has been derived assuming that the system can exchange heat with its surrounding (RHS). If the system as a whole is thermally isolated (which of course does not mean that in the process of reaching the final state a heat exchange cannot occur between its different subsystems) and the final state of the system is prescribed, the inequality (128) becomes the equality: the work done on the MWS in the process in which the system reaches a given final state equals simply the total change of the energy of the system, of the volume source (RVS) and of the matter source (RMS). If however, as in the previously discussed situation, one is interested in a system's change specified only by the final constraints (and not by a concrete final state), the work $\bar{W} = -\Delta(U + p_0V - \mu_0 n)$ done on MWS is maximal when the entropy of the system[91] does not increase; this generalizes the previous considerations by allowing for a change of the system's total volume and matter content due to their exchange with the surrounding (RVS and RMS). If the matter content of the system changes only as a result of chemical reactions (there is no exchange of matter with the surrounding) the work done on MWS is just $\bar{W} = -\Delta(U + p_0V)$ and again it is maximal for that final state which corresponds to the least entropy increase - chemical reactions are nonequilibrium processes and as such always increase entropy. Finally, if the initial and final pressures of all parts of the system are the same and equal to the pressure p_0 of the surrounding (e.g. the chemical reactions in the system occur under the pressure of the surrounding), then the work $\bar{W}$ is given by the change of enthalpy of the system[92]

$$\bar{W} = -\Delta(U + p\,V) = -\Delta H\,. \tag{132}$$

- If the system exchanges with the surrounding only heat (chemical reactions within the system are allowed to occur within the system), and the final total volume of all its parts is identical with the initial one,[93] then $\bar{W} \leq -\Delta(U - T_0S)$. If in addition the initial and final temperatures of all parts of the system are the same and equal to the temperature of RHS (of the heat reservoir with which the system may exchange heat in the process), then, identifying T_0 with T, $\bar{W} \leq -\Delta(U - TS)$ and

  $$\bar{W}_{\rm max} = -\Delta(U - TS) = -\Delta F\,. \tag{133}$$

  In this case the maximal work is given by the change of the Helmholtz free energy of the system.

- If the system exchanges heat and its volume changes but there is no matter exchange with RMS, (chemical reactions within the system are, of course, again allowed to occur within the system), $\bar{W} \leq -\Delta(U - T_0S + p_0V)$. If in addition in the initial and final states all parts of the system are in thermal and mechanical equilibrium with the surrounding ($T = T_0$ and $p = p_0$) then $\bar{W} \leq -\Delta(U - TS + pV)$ and

  $$\bar{W}_{\rm max} = -\Delta(U - TS + pV) = -\Delta G\,. \tag{134}$$

  It is the change of the Gibbs function of the system which in this situation determines the maximal work.

---

[91] More precisely of the supersystem consisting of the system itself, RVS and RMS, but since the entropies of the RVS and RMS are assumed to be invariable, the stated condition pertains only to the system.

[92] Recall that in Lecture II we have argued that $-\Delta H$ is the heat released in the chemical reactions occurring at constant pressure; here - because of the general setting we consider - this heat is assumed to be converted into work within the considered system. (After all, $\bar{W}$ was defined as a change of the total energy of the system and its surrounding).

[93] This can be relaxed if the pressure p_0 of the surrounding vanishes; the change of the total volume of the system can be then assumed to be associated with the work done on the RWS.

To illustrate these results let us consider one simple example. Suppose n moles of a perfect gas at temperature T enclosed in the volume V are given. What maximal work can be obtained by cooling it down to the temperature $T_0 < T$ of the surrounding not changing the gas volume V? As the final volume of the gas is to be the same as the initial one (during the process of cooling down it may vary), the answer to the question is given by

$$\bar{W}_{\text{max}} = -\Delta(U - T_0 S)\,.$$

Using the formulae (which by now should already be well-known!)

$$U(T,V) = C_V T + \text{const.}, \qquad S(T,V) = C_V \ln T + nR\ln(V/n) + \text{const.},$$

(for simplicity constant heat capacity of the gas has been assumed) we readily get the answer

$$\bar{W}_{\text{max}} = C_V(T - T_0) - T_0 C_V \ln\frac{T}{T_0}\,.$$

The process allowing to extract this work can of course be realized with the help of an infinitesimal Carnot engine.[94] One can also ask the question what maximal work can be obtained by cooling the gas to the temperature T_0 of the surrounding keeping it at constant pressure p_0 (or, more generally, ensuring that the initial and final gas pressures are equal p_0), if it initially had the temperature T. The answer is in this case given by the formula $\bar{W}_{\text{max}} = -\Delta(U - T_0 S + p_0 V)$; one now has to consider also the volume changes and compute the entropy change with temperature at constant pressure

$$\Delta V = \frac{nR}{p_0}\,(T_0 - T)\,, \qquad \Delta S = C_p \ln\frac{T_0}{T}\,,$$

and

$$\bar{W}_{\text{max}} = C_V(T - T_0) - T_0 C_p \ln\frac{T}{T_0} - nR\,(T_0 - T) = C_p(T - T_0) - T_0 C_p \ln\frac{T}{T_0}\,.$$

Equilibrium state of a system open to its surrounding and the thermodynamical stability conditions

The results obtained above can be used to formulate the condition determining equilibrium states of systems open to their surrounding (represented by RHS, RVS and RMS), that is when their parameters other than energy and the deformative ones might be directly controlled, generalizing thereby the Callenian maximum entropy principle (which applies when fixed, i.e. directly controlled, values have the deformative parameters characterizing the system as well as its total energy) and to discuss stability conditions which thermodynamical systems should satisfy in typical situations. Some exceptions will be also mentioned.

---

[94] The reader is invited to check it by applying the reasoning similar to the one used on p. 102.

Consider a thermodynamical system which owing to the internal constraints is in equilibrium within itself (various parts of the system are kept in equilibrium by suitable walls). The external constraints separate it from the surrounding represented by a heat reservoir at the temperature T_0, a volume source at p_0 and a reservoir of matter at μ_0. Without the external constraints the system might not be in equilibrium with the surrounding. If it is possible to change the constraints (external and/or internal; in particular, by removing the walls separating it from RHS, RVS and RMS), to which the system is subjected, so that its availability $A = U - T_0 S + p_0 V - \mu_0 n$ decreases, a useful (i.e. positive) work could in principle be extracted from the system through an appropriate process (the formula (128) gives only an upper bound on the work which can be obtained; the system may well attain the new equilibrium state spontaneously without delivering any useful work). If this is not possible, one may say that the system's state is stable: any change of constraints requires supplying to it a positive work from outside (the minimal such work needed is of course equal $W_{\rm min} = +\Delta A > 0$); the constraints separating the system from its surrounding (the RHS, RVS and RMS) can be therefore removed. It follows that in a given surrounding represented by T_0, p_0 and μ_0, it is the minimum of A (over the set all possible virtual equilibrium states that can be fabricated by applying to the system constraints stronger than the ones it is actually subjected to, when it is in equilibrium with its surrounding) which determines the stable equilibrium state of the system. This is the announced generalization to systems open to their surrounding of the Callenian maximum entropy principle (pertaining to systems the energy and the deformative parameters of which are controlled; in particular pertaining to completely isolated systems).

To see what conditions in most typical situations must be satisfied if the system is to be in equilibrium with its surrounding (at T_0 and p_0), we assume (although this needs not always be so - see the example of the equilibrium of a liquid droplet with its vapour discussed in classes and by Pippard) that in equilibrium all parts of the considered system have the same temperature T and the same pressure p and compute (restricting ourselves for a moment to systems which cannot exchange matter with the surrounding and taking the total entropy and the total volume of the system for the parameters which are varied independently) the change of the availability corresponding to arbitrary departures ΔS and ΔV of the system's entropy and volume from the (supposed) equilibrium state[95] (the derivatives are taken at equilibrium values of S and V):

$$A = A^{\rm eq} + \left[\left(\frac{\partial U}{\partial S}\right)_{V,\dots} - T_0\right]\Delta S + \left[\left(\frac{\partial U}{\partial V}\right)_{S,\dots} + p_0\right]\Delta V$$
$$+ \frac{1}{2}\frac{\partial^2 U}{\partial S^2}(\Delta S)^2 + \frac{1}{2}\frac{\partial^2 U}{\partial V^2}(\Delta V)^2 + \frac{\partial^2 U}{\partial V\,\partial S}\Delta V\,\Delta S + \dots \qquad (135)$$

---

[95] The changes ΔS and ΔV need not encompass all possible departures from the system's equilibrium state - there may be also departures corresponding, say, to changes of the numbers $n_1, \dots, n_r$ of moles of matter constituents if chemical reactions can be allowed by changes of system's internal constraints; A has to be minimized with respect to these other departures in the next step.

It follows that the minimum of A (i.e. $\Delta A > 0$ for any departures ΔS and ΔV) is realized by the state in which the temperature T and pressure p (assumed here to be uniform throughout the system) are equal to the temperature T_0 and pressure p_0 of the surrounding, that is, when $T = T_0$ and $p = p_0$. If the considered system is homogeneous this is clear: if $T = T_0$ and $p = p_0$, no useful (positive) work can be extracted from it. However, assuming by the availability A the minimal value requires also that the quadratic form of the second derivatives of A (which translated into the quadratic form of the second derivatives of the internal energy) be (strictly) positive definite. Applying the method of minors (who attended my Math II classes should know what this is) the following conditions are obtained:

$$\left(\frac{\partial^2 U}{\partial S^2}\right)_{V,\dots} = \left(\frac{\partial T}{\partial S}\right)_{V,\dots} = \frac{T}{C_V} > 0\,,$$
$$\left(\frac{\partial^2 U}{\partial V^2}\right)_{S,\dots} = -\left(\frac{\partial p}{\partial V}\right)_{S,\dots} = \frac{1}{V k_S} > 0\,,$$

and (writing $(\partial^2 U/\partial S \partial V)$ in one corner of the matrix of the second derivatives as $-(\partial p/\partial S)_V$ and as $(\partial T/\partial V)_S$ in the other corner)

$$-\left(\frac{\partial p}{\partial V}\right)_S \left(\frac{\partial T}{\partial S}\right)_V + \left(\frac{\partial p}{\partial S}\right)_V \left(\frac{\partial T}{\partial V}\right)_S \equiv \frac{\partial(p,T)}{\partial(S,V)} = \frac{\partial(p,T)}{\partial(V,T)}\frac{\partial(V,T)}{\partial(S,V)} = \frac{T}{V C_V k_T} > 0\,.$$

Thus the stability of the system requires (in the considered case - all parts of the system having the same pressure and temperature) strict positivity of its heat capacity C_V at constant volume and of its adiathermal, k_S, as well as isothermal, k_T, compressibilities.[96] Positivity (or at least nonnegativity of) of k_T is obviously required by the mechanical stability of the system (were it negative, the system would spontaneously compress itself at constant temperature); for similar reason positive should also be the coefficient k_S. Thus under normal circumstances the availability takes its extremal value at $T = T_0$ and $p = p_0$, as expected, and stability (that the extremum is a minimum) is ensured by positive (and not infinite) values of C_V, k_S and k_T. The positivity of these coefficients, necessary for stability, implies, in turn, concavity (downwards or upwards) of various

---

[96] The conditions formulated here can be also given another interpretation: one can consider a small part of an isolated homogeneous system which is in thermal and mechanical equilibrium at the temperature T_0 and pressure p_0; with respect to the small part considered the rest of the system plays precisely the role of the surrounding at T_0 and p_0. This interpretation (encountered in many sources) is, however, only a special case of the much more general situation considered here: we do not assume that the system is homogeneous - it may consist of several parts (subsystems) - but only that in equilibrium all its parts have the same temperature T and the same pressure p (also this can be relaxed - one has then to apply the stability condition in its most general form $\Delta A > 0$); furthermore, as has already been mentioned, departures from the equilibrium may also be due to departures of other variables (other that the system's total entropy S and total volume V) from their equilibrium values - these departures depend on the nature of the system and must, therefore, be considered separately in each particular case (see the examples below).

thermodynamic functions (potentials) characterizing the system. For instance, since

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_V, \qquad \frac{1}{k_T V} = -\left(\frac{\partial p}{\partial V}\right)_T = \left(\frac{\partial^2 F}{\partial V^2}\right)_T,$$

stability implies the concavity upwards of the Helmholz free energy as a function of the temperature T and its concavity downwards as a function of the volume V. Similarly concave upwards, both as a function of the temperature T and pressure p must be the Gibbs function $G(T, p)$. Similar stability conditions can be derived for other simple thermodynamic systems like wires or rubber bands subjected to stretching (in these cases they amount to the inequalities - see Lecture II for definitions of the variables - $C_L > 0$, $(\partial K/\partial L)_S > 0$ and $(\partial K/\partial L)_T > 0$). These also imply definite concavity properties of the corresponding thermodynamic functions of these systems.

There are however well known situations in which some of the derived inequalities become equalities. One such situation is when k_T is infinite, while C_V and k_S are finite.[97] There is then a direction in the $(\Delta S, \Delta V)$ space in which the second order term in the expansion of A around $A^{\rm eq}$ does not grow. This occurs e.g. at the critical point of the liquid-vapour system; one then shows that $(\partial^2 p/\partial V^2)_T$ must vanish while $(\partial^3 p/\partial V^3)_T$ must be negative in order that A has a minimum at $T = T_0$ and $p = p_0$ (equal to the critical values). Indeed, vanishing of k_T^{-1}, i.e. of the determinant of the matrix of the second derivatives, means that (at $T = T_0$, $p = p_0$) the expansion (135) of A can be written in the form

$$A = A^{\rm eq} + \frac{1}{2}\left(\frac{\partial S}{\partial T}\right)_V \left[\left(\frac{\partial T}{\partial S}\right)_V \Delta S + \left(\frac{\partial T}{\partial V}\right)_S \Delta V\right]^2 + \ldots \equiv A^{\rm eq} + \frac{C_V}{2T}(\Delta T)^2 + \ldots,$$

which shows that the direction in question in the $(\Delta S, \Delta V)$ space corresponds to $\Delta T = 0$. Since A treated now as a function of T and V is at constant $T = T_0$ equal to $F(T, V) + p_0 V$, the stability (i.e. the minimum of A at $A^{\rm eq}$ at $p = p_0$) requires, when $(\partial^2 F/\partial V^2)_T = -(\partial p/\partial V)_T = 0$, vanishing of $(\partial^3 F/\partial V^3)_T = -(\partial^2 p/\partial V^2)_T$ and strict positivity of $(\partial^4 F/\partial V^4)_T = -(\partial^3 p/\partial V^3)_T$.

Another possibility would be that all inequalities become equalities (all the three, C_V, k_S and k_T are infinite). As usually with the conditions for a minimum, the third order terms in the expansion of A would then have to vanish and the fourth order terms, the tetra-form[98] of the departures ΔS and ΔV would have to be positive definite. Detailed analysis of this case (Pippard refers here to the Landau and Lifshitz statistical physics textbook) shows that there would be then more conditions than could simultaneously be satisfied, so the conclusion is that C_V, k_S and k_S can never become simultaneously infinite

---

[97] Recall the home Problems in which one shows that $C_p = C_V + TV\alpha_p^2/k_T$ and $k_T = (C_p/C_V)k_S$. From these relations it follows that $C_p \geq C_V$ (the equality requires either $\alpha_p \equiv (1/V)(\partial V/\partial T)_p 0$ or infinite k_T) and, consequently, $k_T \geq k_S$.

[98] Tetra-form like tetradrachm - a Greek coin, an artefact from antiquity - or tetrarchy - a system (in effect not successful, as the subsequent history has shown) of governing the Roman empire established by Diokletianus.

if A can be expanded in powers of the departures ΔS and ΔV around the equilibrium state.

It may also happen that A cannot be expanded in the power series. The typical situation is a (nonhomogeneous) system consisting of two phases α and β in equilibrium with the surrounding at T_0 ad p_0 which are such that both phases may coexist with an arbitrary proportion n_α/n_β ($n_\alpha + n_\beta = n$) of the matter in each of the two phases. In such a case $k_T = \infty$ because changing the volume does not change the pressure - only the proportion n_α/n_β is thereby altered. So A has a valley of equal minima along a direction in the (S, V) plane and begins to raise only at the opposite ends of this valley corresponding to the volumes in which one of the two phases completely disappears. As long as both phases are present, the system's equilibrium is neutral (see the discussion in Lecture I). It is amusing to show that considering this valley allows to derive the Clapeyron-Clausius equation for the temperature dependence of the pressure along the coexistence curve.[99] The valley must be in the null direction of the quadratic form of the second derivatives in (135):

$$\begin{pmatrix} (\partial T/\partial S)_V & (\partial T/\partial V)_S \\ (\partial T/\partial V)_S & -(\partial p/\partial V)_S \end{pmatrix} \begin{pmatrix} \Delta S \\ \Delta V \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{136}$$

From the first line of this equality (the second line give a linearly dependent equation - this is ensured by vanishing of the determinant of the above matrix, that is, by the infinite value of k_T) one gets that the valley direction is such that

$$\frac{\Delta S}{\Delta V} = -\frac{(\partial T/\partial V)_S}{(\partial T/\partial S)_V} = \left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V , \tag{137}$$

(on the right hand side first the “shocking” relation has been used and then the well-known Maxwell identity). Along the valley however, the entropy and the volume changes are given by

$$\Delta S = s_\alpha \Delta n_\alpha + s_\beta \Delta n_\beta \, , \qquad \Delta V = v_\alpha \Delta n_\alpha + v_\beta \Delta n_\beta \, ,$$

but since $n_\alpha + n_\beta = n$ is fixed, $\Delta n_\alpha = -\Delta n_\beta$ and therefore (137) takes the form[100]

$$\frac{dp}{dT} = \frac{s_\alpha - s_\beta}{v_\alpha - v_\beta} , \tag{138}$$

which is just the Clapeyron-Clausius equation (to which we will return in the last Lecture devoted to thermodynamics). In most cases neither the numerator (i.e. ΔS) nor the denominator (i.e. ΔV) on the right hand side of (138) vanishes, which means that neither $(\partial^2 U/\partial S^2)_V$ nor $(\partial^2 U/\partial V^2)_S$ vanish, that is, C_V and k_S are both finite (though k_T, C_p

---

[99] This equation will be rederived with the help of a more conventional and more general reasoning in Lecture IX.

[100] Since along the valley the pressure does not depend on the volume, the derivative $(\partial p/\partial T)_V$ acquires the meaning of the derivative dp/dT along the coexistence curve.

and α_p are infinite). On the melting curve of ${}_2\mathrm{He}^3$ (the lighter isotope of Helium) there is, however, a point at which $dp/dT = 0$ which means that the valley lies in the $(0, \Delta V)$ direction; this in turn implies (since the matrix in (136) must vanish on this vector) that $(\partial^2 U/\partial V^2)_S = 0$, that is, that $k_S = \infty$. Similarly a point at which $dT/dp = 0$ would probably exist on the melting curve of ice were it not for the transformation of ordinary ice into its another phase (H_2O has many different phases), which in the analogous way would mean vanishing of $(\partial^2 U/\partial S^2)_V = 0$, that is $C_V = \infty$. This shows that the finitness of C_V and k_S are not absolute thermodynamic requirements (but exceptions are rare).

Special cases of the general stability criterion $\Delta A > 0$ (determining in the variational way the equilibrium states of thermodynamic systems in different conditions) formulated above deserve consideration.

- If the system is isolated thermally and with respect to the matter transfer (no contact with RHS and RMS; chemical reactions inside it can occur), $U + p_0 V$ stays constant[101] and the condition of minimal value of the availability $A = U + p_0 V - T_0 S$ is equivalent to the condition of maximum entropy discussed before.

- If the volume of the system in contact with RHS at T_0 is fixed (there is no contact with RVM), then $p - p_0$ in the expansion (135) is indeterminate (the equilibrium pressure in the system is unrelated to that of the surrounding). The minimum of A is at $T = T_0$ and the other parameters specifying the state of the system are determined by minimizing the free Helmholtz energy $F(T, V)$ over the set of all virtual states realizable with the help of stronger constraints keeping constant V and constant $T = T_0$.

- If the system is in thermal and mechanical contact with its surrounding (i.e. with RHS and TVS) at T_0 and p_0, and *if* all its parts have in equilibrium the same temperature T and the same pressure p, then $T = T_0$ and $p = p_0$ and the values of the remaining parameters specifying its states are determined by minimizing the Gibbs function $G = U - TS + p\,V$ at fixed $T = T_0$ and $p = p_0$ over the set of all virtual states realizable with the help of stronger constraints compatible with the fixed temperature and presure (equal T_0 and p_0) of all parts of the system.

As an example of applications of the stability conditions discussed above, we can consider the equilibrium of the system consisting of n moles of a liquid and its vapour, under different conditions. For simplicity we can assume that there is only one matter component. Let first the vessel containing the system be open to a constant external pressure p_0 and held at a temperature T_0 by its thermal contact with the surrounding (playing the role of the heat-bath). The availability of the system corresponding to virtual states in which the transfer of matter between the liquid and the vapour is blocked by a

---

[101] I.e. one considers only such virtual states; if the system is isolated both thermally and mechanically, U and V are separately conserved; if it is only thermally isolated but in mechanical contact with RVS, the combination $U + p_0 V$ is conserved, because $U + U_{\mathrm{RVS}}$ is.

suitable wall is

$$A = n_{\rm l}(u_{\rm l} - T_0 s_{\rm l} + p_0 v_{\rm l}) + n_{\rm v}(u_{\rm v} - T_0 s_{\rm v} + p_0 v_{\rm v})\,, \tag{139}$$

where $n_{\rm l}$ ($n_{\rm v}$) is the number of moles of the liquid (vapour) in the vessel. Obviously, $n_{\rm l} + n_{\rm v} = n$ and all quantities $u_{\rm l,v}$ etc., are computed at T_0 and p_0 (they are therefore fixed) because we expect that in equilibrium both subsystems (the liquid and the vapour) should have the same temperatures and pressures[102] and from the previous considerations we know that these must be equal to T_0 and p_0, respectively. The equilibrium values $n_{\rm l}^{\rm eq}$ and $n_{\rm v}^{\rm eq}$ (also satisfying the relation $n_{\rm l}^{\rm eq} + n_{\rm v}^{\rm eq} = n$) should minimize A (or equivalently, the Gibbs function, $G(T_0, p_0, n_{\rm l}, n_{\rm v})$ at constant T_0 and p_0). That is, if both phases are present, δA should vanish at $n_{\rm l} = n_{\rm l}^{\rm eq}$. From this we find the equilibrium condition

$$u_{\rm l}(T_0, p_0) - T_0 s_{\rm l}(T_0, p_0) + p_0 v_{\rm l}(T_0, p_0) = u_{\rm v}(T_0, p_0) - T_0 s_{\rm v}(T_0, p_0) + p_0 v_{\rm v}(T_0, p_0)\,,$$

that is, equilibrium is possible only if T_0 and p_0 are such that $g_{\rm l}(T_0, p_0) = g_{\rm v}(T_0, p_0)$ or, as we will see shortly, $\mu_{\rm l}(T_0, p_0) = \mu_{\rm v}(T_0, p_0)$. One should also notice that in this case the second and all higher derivatives of A with respect to $n_{\rm l}$ vanish identically - the equilibrium is of the neutral nature (the value of $n_{\rm l}$ is not fixed by the equilibrium condition); this has already been discussed in this Lecture. If T_0 and p_0 are not such that $g_{\rm l}(T_0, p_0) = g_{\rm v}(T_0, p_0)$, only one phase - the one of lower value of the molar Gibbs function - is present.

As the second situation we consider the same mixture of n moles of a liquid and its vapour but now as an isolated system of fixed total energy $U = n_{\rm l}u_{\rm l} + n_{\rm v}u_{\rm v}$ and fixed volume $V = n_{\rm l}v_{\rm l} + n_{\rm v}v_{\rm v}$. Minimization of the system's availability (139) reduces now to maximizing its total entropy $S = n_{\rm l}s_{\rm l} + n_{\rm v}s_{\rm v}$, because U and V (and n) are fixed. The external pressure p_0 and the temperature T_0 do not play any role here (the system is isolated from the external pressure and from the heat bath). However to seek the extremum of S respecting the constraints one usually uses the standard method[103] and equates to zero the variation of the the auxiliary function

$$\delta(U - TS + pV) = 0\,,$$

in which the parameters T and p are the Lagrange multipliers,[104] treating the variables $n_{\rm l}$ and, say, $u_{\rm l}$, $v_{\rm l}$, $u_{\rm v}$, $v_{\rm v}$ as independent. The virtual equilibrium states which should be

---

[102] The pressures of the two phases may be nonequal if the presence of the surface separating them is taken into account.

[103] Of course, were we given the explicit forms of $s_{\rm l}(u_{\rm l}, v_{\rm l})$ and $s_{\rm v}(u_{\rm v}, v_{\rm v})$, we could write down the total entropy $S = n_{\rm l}s_{\rm l}(u_{\rm l}, v_{\rm l}) + n_{\rm v}s_{\rm v}(u_{\rm v}, v_{\rm v})$ as an explicit function $S = S(U, V, n, u_{\rm l}, v_{\rm l}, n_{\rm l})$, say, and maximize it directly treating $0 \leq u_{\rm l} \leq U/n$, $0 \leq v_{\rm l} \leq V/n$ and $0 \leq n_{\rm l} \leq n$ as three independent variables. If both phases are present in equilibrium (that is, if the derivatives of $S(U, V, n, u_{\rm l}, v_{\rm l}, n_{\rm l})$ with respect to $u_{\rm l}$, $v_{\rm l}$ and $n_{\rm l}$ vanish within the specified ranges), they both should have the same T and p and, therefore, all the molar quantities $u_{\rm l}$, $v_{\rm l}$, $s_{\rm l}$ and $u_{\rm v}$, $v_{\rm v}$, $s_{\rm v}$ can be parametrized by the temperature T and pressure p obtainable as $\partial s_{\rm l}/\partial u_{\rm l} = 1/T = \partial s_{\rm v}/\partial u_{\rm v}$ and $\partial s_{\rm l}/\partial v_{\rm l} = p/T = \partial s_{\rm v}/\partial v_{\rm v}$; the equilibrium values $u_{\rm l}$, $v_{\rm l}$, $u_{\rm v}$, $v_{\rm v}$ can be thus eventually traded for the variables T and p.

[104] In this simple case there is no need to introduce the Lagrange multiplier associated with the condition

considered in this situation may correspond to different partitions of the total internal energy U, the total volume V and numbers of moles (and, hence, to different entropies) between the two phases. Therefore the above variation can be written as

$$\begin{aligned} &n_{\rm l}[\delta u_{\rm l} - T\delta s_{\rm l}(u_{\rm l}, v_{\rm l}) + p\,\delta v_{\rm l}] + n_{\rm v}[\delta u_{\rm v} - T\delta s_{\rm v}(u_{\rm v}, v_{\rm v}) + p\,\delta v_{\rm v}] \\ &\quad + \delta n_{\rm l}[u_{\rm l} - Ts_{\rm l}(u_{\rm l}, v_{\rm l}) + p\,v_{\rm l}] + \delta n_{\rm v}[u_{\rm v} - Ts_{\rm v}(u_{\rm v}, v_{\rm v}) + p\,v_{\rm v}] = 0\,. \end{aligned}$$

We know however, that the two Lagrange multipliers T and p can be chosen in such a way as to kill the coefficients of $n_{\rm l}$ and $n_{\rm v}$. This is possible, because $\delta u_{\rm l}$, $\delta v_{\rm l}$ and $\delta s_{\rm l}$ are changes of the parameters between two equilibrium states (the real one and a virtual one) of one mole of the isolated liquid and, as has been stated at the end of Lecture V, these changes are always correlated in this way with T and p being the temperature and pressure of the system; the same applies also to the changes $\delta u_{\rm v}$, $\delta v_{\rm v}$ and $\delta s_{\rm v}$. Thus in this way, the Lagrange multipliers acquire the proper meaning of the system's temperature and pressure and the equilibrium condition is

$$\delta n_{\rm l}(u_{\rm l} - Ts_{\rm l} + pv_{\rm l})|_{T,p} + \delta n_{\rm v}(u_{\rm v} - Ts_{\rm v} + pv_{\rm v})|_{T,p} = 0\,,$$

that is, because $\delta n_{\rm l} = -\delta n_{\rm v}$,

$$g_{\rm l}(T,p) = g_{\rm v}(T,p)\,, \tag{140}$$

which of course must be solved together with the conditions

$$n_{\rm l}u_{\rm l}(T,p) + n_{\rm v}u_{\rm v}(T,p) = U\,, \qquad n_{\rm l}v_{\rm l}(T,p) + n_{\rm v}v_{\rm v}(T,p) = V\,, \tag{141}$$

to yield the equlibrium parameters T, p, $n_{\rm l}$ and $n_{\rm v} = n - n_{\rm l}$. The condition (140) is the same as $\mu_{\rm l}(T,p) = \mu_{\rm v}(T,p)$ (because there is only a single component). Of course, now the matrix of the second derivatives is not identically zero - and and if the conditions (140) and (141) can be met simultaneously, it is negative definite (at least on the departures of the variables respecting the constraints). If they cannot be met, there is only one phase - the one that has lower g, that is, in this case, that of higher entropy.

Finally on can also consider the equilibrium of the liquid with its vapour in the vessel of fixed volume but held at constant temperature T_0 by a thermal contact with the surrounding. In this case minimizing A reduces, because V is fixed, to minimizing $U-$

$n_{\rm l}+n_{\rm v} = n$ because it can be easily taken into account directly. In more complicated cases (several phases, more matter components) one would have to introduce one additional Lagrange multipliers (which would acquire the meaning of its chemical potential - see Lecture IX) per each component. Maximizing S it would be more appropriate to write $\delta(S - \lambda_1 U - \lambda_2 V) = 0$, as in similar Math II problems, with the Lagrange multipliers λ_1 and λ_2 which (because of the resulting conditions $\partial s_{\rm l}/\partial u_{\rm l} = \lambda_1 = \partial s_{\rm v}/\partial u_{\rm v}$ and $\partial s_{\rm l}/\partial v_{\rm l} = \lambda_2 = \partial s_{\rm v}/\partial v_{\rm v}$ would then acquire the interpretation of $1/T$ and p/T, respectively, but this is clearly equivalent to writing $\delta(U - TS + pV) = 0$. It should be also noted that the written condition can be also interpreted as seeking the minimum of U at constant S (and V) - this illustrates the equivalence of the maximum entropy and minimum internal energy rules which both can be used to determine the equilibrium state of an isolated system.

$T_0 S$, or because the external temperature T_0 is also the temperature of the system, to minimizing[105]

$$F(T_0, V, v_\mathrm{l}, v_\mathrm{v}, n_\mathrm{l}) = n_\mathrm{l} f_\mathrm{l}(T_0, v_\mathrm{l}) + (n - n_\mathrm{l}) f_\mathrm{v}(T_0, v_\mathrm{v}) \, ,$$

with the condition of fixed $V = n_\mathrm{l} v_\mathrm{l} + n_\mathrm{v} v_\mathrm{v}$ which is taken into account by introducing the Lagrange multiplier p, which then acquires the meaning of the internal pressure of the system. Similar reasoning as above then leads to the conditions

$$g_\mathrm{l}(T_0, p) = g_\mathrm{v}(T_0, p) \, , \qquad n_\mathrm{l} v_\mathrm{l}(T_0, p) + n_\mathrm{v} v_\mathrm{v}(T_0, p) = V \, ,$$

which together determine the equilibrium pressure p and the numbers n_l and n_v of moles in the two phases. Also in this case if all the conditions can be met, the extremum of the function $F(T_0, V, v_\mathrm{l}, v_\mathrm{v}, n_\mathrm{l})$ is its true minimum (with respect to departures δv_l, δv_v and δn_l respecting the constraints).

That in all the three cases the condition of stability (determining the equilibrium state) ultimately boils down to the requirement of equality of the molar Gibbs functions of the two phases should not be surprising: if the equilibrium has been reached in the isolated system and corresponds to a common temperature T and a common pressure p of both phases, then it will not be destroyed by placing the system in contact with the surrounding at that temperature and that pressure.

---

[105] Again, if the explicit forms of $f_\mathrm{l}(T_0, v_\mathrm{l})$ and $f_\mathrm{v}(T_0, v_\mathrm{v})$ were known, one could directly minimize $F(T_0, V, n, v_\mathrm{l}, n_\mathrm{l}) \equiv n_\mathrm{l} f_\mathrm{l}(T_0, v_\mathrm{l}) + n_\mathrm{v} f_\mathrm{v}(T_0, v_\mathrm{v})$ with respect to the variables v_l and n_l after eliminating n_v and v_v using the conditions $n = n_\mathrm{l} + n_\mathrm{v}$ and $V = n_\mathrm{l} v_\mathrm{l} + n_\mathrm{v} v_\mathrm{v}$. The equilibrium pressure could be then determined as, say, $-\partial f_\mathrm{l}/\partial v_\mathrm{l}$ (computed at the extremum) because in equilibrium the pressure must be the same for both phases.

LECTURE VIII (TMD)

Fundamental relations in the representations of the F, G and H functions
The considerations carried out in the preceding Lecture showed that in specific conditions, when the system is open to its surrounding (is held at constant temperature and/or pressure through the contact with suitable reservoirs representing the surrounding), its equilibrium states are determined by minima (over the set of all possible virtual equilibrium states compatible with the external conditions) of the Helmholz free energy F (if T and V are fixed), or the enthalpy H (if S and p are fixed) or the Gibbs function G (if T and p are fixed). The principles of minimum of F, H or G in these situations replace the Callenian principle of maximal entropy. In this sense all these functions also play the roles of **thermodynamic potentials**. We have also argued that if the entropy S of an isolated system is known as a function of its global (extensive, if the system has this property) parameters U, V, … and n (or $n_1, \dots, n_r$), the thermodynamical information about the system is complete. We now want to argue that if the termodynamic potentials are known as functions of their **natural variables**: F as a function of T, V and n (or $n_1, \dots, n_r$), H as a function of S, p and n or of G as a function of T, p and n, the thermodnamical information about the system is also complete. The distinguished role of the particular thermodynamical potential in given external conditions stems precisely from the fact that its natural variables are just those which in the given situation are directly controlled (by the environment).

From the mathematical point of view the operations which one does passing from the internal energy U known as a function of S, V and n to the other potentials in their natural variables is called **Legendre transformation**. This transformation in its simplest form is precisely the way of going over from a given function $f(x)$ to another function $g(p)$, the argument p of which is the derivative of f, without loosing the information about the form of f (only coding it differently). Since the Legendre transformation is used in many places in physics, it is appropriate to discuss it here in the general way.

Let f represent a physical quantity which is theoretically given as a convex[106] function (upwards or downwards) function $f(x)$ of some variable x. Suppose, however, that directly controlled experimentally is not the variable x itself but the derivative p of f with respect to x. One could then try to invert the relation

$$p(x) = df/dx\,,$$

to get $x = x(p)$ and to switch to the function

$$\tilde{f}(p) \equiv f(x(p))\,. \tag{142}$$

which would represent the same physical quantity (because of this a physicist would just write $f(p)$). This is precisely what one does expressing U in terms of T instead of

---

[106] The generalization of the Legendre transform to nonconvex functions is called the Fenchel transform and is defined as $g(p) = \sup_x(f(x) - p\,x)$ and plays the important role in the theory of phase transitions.

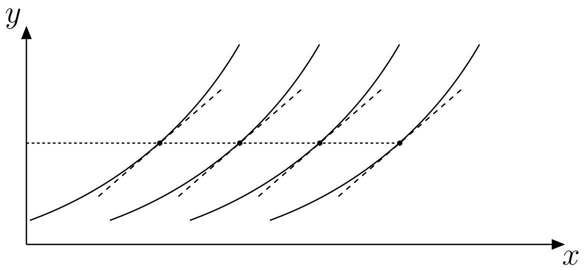

Figure 17: A family of functions $y = f_a(x)$ which all yield the same function $\tilde{f}(x(p))$, where $p = p(x)$ is the slope of the tangent to $f_a(x)$ at x.

expressing it in terms of S. But switching to $\tilde{f}(p)$ one looses some information about the form of $f(x)$: the functions $f(x)$ and $f_a(x) \equiv f(x - a)$ with an arbitrary shift constant a lead to the same $\tilde{f}(p)$ (see Figure 17).

However the set of lines tangent at every point to a given convex curve on the plane (x, y) determines this curve uniquely: geometrically the plot of $y = f(x)$ is just the envelope of the set these tangent lines. In turn, every tangent line is uniquely determined by its slope p (the variable we want to play with) and the value g of the intersection of the tangent with the y-axis. It suffices therefore to give g as a function of the slope (defining in this way the family of the tangent lines) to retain the complete information about the form of the original function $f(x)$. To this end we consider a point x_0 and write down the equation of the tangent to $f(x)$ at this point (see Figure 18):

$$y = p\,x + f(x_0) - p\,x_0\,.$$

The value g of the intersection of this tangent with the y-axis is therefore equal (renaming now x_0 to x)

$$g = f(x) - p\,x\,.$$

If now x is written as $x = x(p)$ inverting the relation (142) we will get the function

$$g(p) = f(x(p)) - p\,x(p)\,,$$

which encodes in it the same information about the dependence of the quantity f on the variable x as does the original function $f(x)$, but is expressed through the variable which is easily controlled experimentally. That the combination $f(x) - px$ is indeed naturally a function of the variable p can be also seen by considering its differential

$$dg \equiv d(f(x) - p\,x) = df - d(p\,x) = \frac{df}{dx}\,dx - p\,dx - x\,dp = -x(p)\,dp\,.$$

Thus, the change of g depends only on the change dp of the argument p and not on dx; the function $g(p)$ can be therefore reconstructed from its differential dg by the step by step integration given its value $g(p_0)$ at some point p_0.

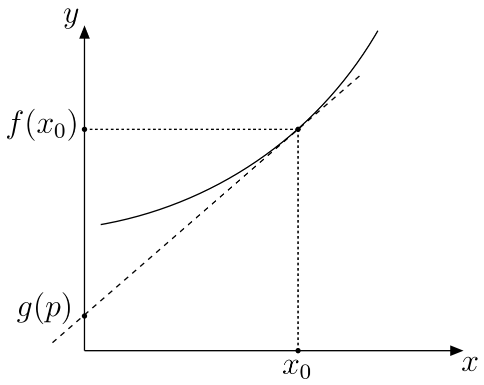

Figure 18: Geometrical construction of the Legendre transform g value $g(p)$ corresponding to the value $f(x_0)$ of a convex function f.

Of course a function of several variables, $f = f(x_1, \dots, x_n)$ can be Legendre transformed in an arbitrary number of its variables to obtain, say,

$$g(p_1, \dots, p_k, x_{k+1}, \dots, x_n) = f(x_1, \dots, x_n) - \sum_{j=1}^{k} p_j x_j \,, \tag{143}$$

where the variables $x_1, \dots, x_k$ should be expressed in terms of the variables $p_1, \dots, p_k$ (and $x_{k+1}, \dots, x_n$) by inverting with respect to them the k relations

$$p_i = \frac{\partial f}{\partial x_i} \,. \tag{144}$$

Students associate the Legendre transform primarily with classical mechanics in which it is used to pass from the Lagrange to Hamilton's formulations of the equations of motion (on this other formulation of mechanics is based the transition - traditionally called quantization - to the quantum theory; it is also fundamental to the formulation of classical statistical mechanics, which we will discuss in due course), i.e. to replace the generalized variables q^l and the generalized velocities $\dot{q}^l$, $l = 1, \dots, n$ by the variables q^l and the canonical momenta $p_l = \partial L/\partial q^l$. (Because of this, students have the tendency to think that the Legendre[107] transform has something to do with the fact that in mechanics variables come in pairs q^i and $\dot{q}^l$. The derivation above clearly shows this is not so.) The Hamiltonian $H(p_1, \dots, p_n, q^1, \dots, q^n)$ is just the (minus, in order that it has - in most cases - the interpretation of mechanical energy) Legendre transform of the Lagrangian function $L = L(q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n)$ in n its last variables

$$H(p_1, \dots, p_n, q^1, \dots, q^n) = \sum_{j=1}^{n} p_j \dot{q}^j - L(q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n) \,.$$

It is also instructive to illustrate the working of the Legendre transform on another simple physical example taken from electrostatics. Let us determine the force by which

---

[107] Legendre, Lagrange, Laplace, Lavoisier, Lebesque - who can distinguish all these French L-masters?

the two plates of a capacitor attract one another. We remember that the voltage φ between the plates is related to the charge $\mathcal{Q}$ (on one of the plates) and the capacity $\mathscr{C}$ by $\mathcal{Q} = \mathscr{C}\varphi$. If the capacitor is being charged by successively bringing onto its plate infinitesimal portions $d\mathcal{Q}$ of charge (e.g. in the flat capacitor by transporting successively and reversibly - in the sense that an external force only counterbalances the electric forces acting on the transported charge - portions $d\mathcal{Q}$ from one plate to the other one) the work $đW$ done on the capacitor by external forces is equal $đW = \varphi\, dQ$. Charging the capacitor with the charge $\mathcal{Q}$ in this way requires doing on this system the work

$$\int_0^{\mathcal{Q}} đW = \int_0^{\mathcal{Q}} \varphi\, d\mathcal{Q} = \int_0^{\mathcal{Q}} \frac{\mathcal{Q}}{\mathscr{C}}\, d\mathcal{Q} = \frac{\mathcal{Q}^2}{2\mathscr{C}} \equiv U\,,$$

which is therefore equal to the energy of the charged capacitor (its internal energy, if we wish to treat the capacitor in the thermodynamic way).

If the capacity $\mathscr{C}$ of the capacitor is altered as a result of the action of an external force F (when the application of the external force changes the distance between the plates - in the language of thermodynamics - reversibly), the work done by this external force is just the change of the internal energy U of the capacitor, provided the capacitor is an isolated system, that is, has a fixed charge $\mathcal{Q}$ (we do not consider here the heat capacity of the plates). This allows to find the force by which the plates attract each other. (This in Feynman's Lectures on Physics is called the principle of virtual works but in fact this is precisely the same as applying 1TMDL to isolated systems - in thermodynamics adiathermally isolated systems; here we neglect possible thermal effects but the capacitor could be not isolated also because of being connected to a battery, so we need to exclude this possibility and this is just the condition of the constancy of $\mathcal{Q}$.) Let us write this statement in the thermodynamical way

$$dU(\mathcal{Q}, \mathscr{C}) = \left(\frac{\partial U}{\partial \mathcal{Q}}\right)_{\mathscr{C}} d\mathcal{Q} + \left(\frac{\partial U}{\partial \mathscr{C}}\right)_{\mathcal{Q}} d\mathscr{C} \equiv \varphi\, d\mathcal{Q} + \mathscr{F}\, d\mathscr{C}\,,$$

By $\mathscr{F}$ we have denoted here the generalized force related to the change of the capacity at fixed charge $\mathcal{Q}$ of the capacitor. Expressing the capacity $\mathscr{C}$ through the spatial (geometrical) characteristics of the capacitor (the area of its plates, the distance between them) allows to give the generalized force $\mathscr{F}$ the ordinary mechanical meaning. For example if $\mathscr{C}$ changes due to changing the distance z between the plates

$$dU(\mathcal{Q}, \mathscr{C})|_{\mathcal{Q}\,=\,\text{const.}} = \mathscr{F}\, d\mathscr{C} = \mathscr{F}\, \frac{d\mathscr{C}}{dz}\, dz \equiv F_z\, dz\,.$$

In this way we find that the external force which precisely balances the force F_z by which the plates of the flat capacitor of area A and separated by the distance z between the plates (the capacity of such a capacitor is in the SI system equal $\mathscr{C} = \varepsilon_0(A/z)$) attract each other is given by

$$F_z = \left(\frac{\partial U}{\partial \mathscr{C}}\right)_Q \frac{d\mathscr{C}}{dz} = \left(\frac{\partial U}{\partial z}\right)_Q = \frac{Q^2}{2\varepsilon_0 A}\,.$$

The force F_z found in this way cannot depend, of course, on whether the capacitor has a fixed charge $\mathcal{Q}$, or whether it is connected to a battery which keeps it at the voltage φ of such a magnitude that the charge on the plates is equal $\mathcal{Q}$. If we write, however, the energy U as a function $U(\varphi, \mathscr{C})$, expressing simply the charge through the voltage and the capacity and try to define the generalized force $\mathscr{F}$ acting between the capacitor plates at constant voltage φ by the relation

$$dU(\varphi, \mathscr{C})|_{\varphi\,=\,\text{const.}} = \left(\frac{\partial U}{\partial \mathscr{C}}\right)_{\varphi} d\mathscr{C} \equiv \mathscr{F}\, d\mathscr{C} \qquad \text{(incorrect)},$$

we will get the force F_z with the wrong sign. To obtain the right sign one has to perform the Legendre transform, that is, to pass to the function

$$\tilde{U}(\varphi, \mathscr{C}) = U(\mathcal{Q}(\varphi, \mathscr{C}),\, \mathscr{C}) - \varphi\, \mathcal{Q}(\varphi, \mathscr{C})\,,$$

the "free energy", the differential of which is

$$\begin{aligned} d\tilde{U}(\varphi, \mathscr{C}) &= \varphi\, d\mathcal{Q}(\varphi, \mathscr{C}) + \mathscr{F}\, d\mathscr{C} - d(\varphi \mathcal{Q}) \\ &= \left(\frac{\partial \tilde{U}}{\partial \varphi}\right)_{\mathscr{C}} d\varphi + \left(\frac{\partial \tilde{U}}{\partial \mathscr{C}}\right)_{\varphi} d\mathscr{C} \equiv -\mathcal{Q}\, d\varphi + \mathscr{F}\, d\mathscr{C}\,, \end{aligned}$$

As we can now fully control the voltage, we can compute the force F_z as the coefficient in $d\tilde{U}$ of the dz differential at fixed voltage φ

$$\begin{aligned} d\tilde{U}(\varphi, \mathscr{C})\Big|_{\varphi=\text{const.}} &= \left(\frac{\partial}{\partial \mathscr{C}}\left[\frac{1}{2}\mathscr{C}\varphi^2 - \varphi\, \mathcal{Q}(\varphi, \mathscr{C})\right]\right) d\mathscr{C} \\ &= \left(\frac{\partial}{\partial \mathscr{C}}\left[-\frac{1}{2}\mathscr{C}\varphi^2\right]\right) d\mathscr{C} = -\frac{1}{2}\varphi^2 \frac{d\mathscr{C}}{dz}\, dz \equiv F_z\, dz\,. \end{aligned}$$

In this way one gets the right force F_z (with the right sign) but expressed through different variables.

The physical reason for the necessity to pass to the function $\tilde{U}$ is, of course, that the capacitor connected to the battery is not an isolated system and changing its capacity $\mathscr{C}$ by moving its plates entails doing on it some work also by the battery, which has to supply to the capacitor an additional charge (to maintain the voltage unchanged). The energy balance in this case therefore reads

$$F_z\, dz + đW_{\text{bat}} = dU(\varphi, \mathscr{C}) \equiv \frac{1}{2}\varphi^2\, d\mathscr{C}\,.$$

The work done by the battery goes into supplying the additional charge $\varphi d\mathscr{C}$ to the capacitor at the voltage φ; therefore this work equals $đW_{\text{bat}} = \varphi^2 d\mathscr{C}$. Putting this work on the other side of the above equality yields the right force F_z. In the "thermodynamic" approach consisting of applying the principle of virtual works to the function $\tilde{U}$ instead of U, the work done by the battery is already automatically taken into account by controlling the voltage φ. One should notice here the analogy to using the Helmholtz free energy F instead of the internal energy U in computing the volume work done by a system which is kept in equilibrium witha heat bath at fixed temperature T

In thermodynamics with the help of the Legendre transform two families of potentials can be constructed. The first one (more commonly used) is obtained by starting from the fundamental relation (102) in the internal energy representation $U = U(S, V, \dots, n_1, \dots, n_r)$. The three basic potentials of a simple system (a fluid) are the already introduced enthalpy (to simplify the notation n stands for $n_1, \dots, n_r$ and $\mu\, dn$ for $\sum_{j=1}^{r} \mu_j dn_j$)

$$H(S, p, n) = U + p\,V\,, \tag{145}$$
$$dH = T(S, p, n)\, dS + V(S, p, n)\, dp + \mu(S, p, n)\, dn\,, \tag{146}$$

the Helmholtz free energy

$$F(T, V, n) = U - TS\,, \tag{147}$$
$$dF = -S(T, V, n)\, dT - p(T, V, n)\, dV + \mu(T, V, n)\, dn\,, \tag{148}$$

and the Gibbs function (called also the free enthalphy)

$$G(T, p, n) = U - TS + p\,V\,, \tag{149}$$
$$dG = -S(T, p, n)\, dT + V(T, p, n)\, dp + \mu(T, p, n)\, dn\,. \tag{150}$$

One also defines the Grand potential

$$\Omega(T, V, \mu) = U - TS - \mu\, n\,, \tag{151}$$
$$d\Omega = -S(T, V, \mu)\, dT - p(T, V, \mu)\, dV - n(T, V, \mu)\, d\mu\,. \tag{152}$$

In the similar way one defines the functions H, F and G characterizing other simple systems like wires and rubber bands ($-p \to K$, $V \to L$), to films ($-p \to \gamma$, $V \to A$) or to magnetic materials ($-p \to \mathcal{H}_0$ in the normal Gauss system or $-p \to \mu_0\mathcal{H}_0$ in the SI system - but if μ is used to denote the chemical potential, the SI system becomes clearly inconvenient - and $V \to M$). If the system is not simple and has more variables, for example a paramagnetic gas the internal energy of which depends, in addition to the number of moles, on S, V and M, more Legendre transforms can be formed and their names are not codified. One of such potentials has already been used in discussing the connection between piezoelectricity and electrostriction in Lecture V.

As follows from the construction of the Legendre transform, all these potentials, if known as functions of their natural variables (those which are explicitly indicated in the formulae above), contain the complete thermodynamical information about systems to which they pertain. Therefore the relations $H = H(S, p, n)$, $F = F(T, V, n)$ or $G = G(T, p, n)$ can be called fundamental relations in the representations of enthalpy, free energy and Gibbs, respectively. To show this directly it is sufficient to realize that from each of them $U = U(S, V, \dots, n)$ can be obtained by the repeated Legendre transform (the Legendre transform applied twice to a convex function is the identity operation). For instance, knowing $F = F(T, V, n)$ one can write

$$U(T, V, n) = F + TS = F - T\left(\frac{\partial F}{\partial T}\right)_{V,n} \equiv -T^2\left(\frac{\partial}{\partial T}\,\frac{F}{T}\right)_{V,n}, \tag{153}$$

and inverting the relation $S = -(\partial F/\partial T)_{V,n}$ to get $T = T(S, V, n)$ one can obtain $U = (S, V, n)$. Moreover the potentials allow (as has been demonstrated in classes) to obtain Maxwell identities more straightforwardly than does $U = U(S, V, n)$. For instance the identity $(\partial S/\partial V)_{T,n} = (\partial p/\partial T)_{V,n}$ is an immediate consequence of the fact that F is a state function and therefore its mixed second derivatives must be equal. As another illustration of the usefulness of F we can obtain the dependence of the heat capacity C_V on the volume:

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

The distinguished role of the potentials $F(T, V, \dots, n)$, $\Omega(T, V, \dots, \mu)$ and, primarily, $S(U, V, \dots, n)$ stems from the fact that the formalism of the equilibrium statistical physics gives direct prescriptions to determine them on the basis of the microscopic dynamics (classical or quantum) of the considered system when the latter is subjected to specific constraints: the entropy S, if the energy, the amount of matter and all deformative parameters of the system are directly controlled (e.g. if the system is thermally isolated), the free Helmholtz energy $F(T, V, n)$, if directly controlled is the system's temperature, the amount of matter in it as well as the relevant deformative variables (e.g. if a closed system is isolated from the external pressure but remains in thermal contact with a heat bath at temperature T) and the potential $\Omega(T, V, \mu)$ when directly controlled are its temperature T, the relevant deformation parameters and the chemical potential, e.g. when system can exchange matter with a reservoir at the chemical potential μ. (Although the Gibbs potential $G(T, p, n)$, relevant when controlled are temperature, pressure and the amount of matter, can also be obtained directly by considering the appropriate statistical ensemble - the notion to be defined yet, it is less frequently used.)

The other family of thermodynamic potentials (used practically only by a very narrow family of thermodynamics specialists) is constructed taking as the starting point the fundamental relation in the entropy representation (97) the differential of which

$$dS = \frac{1}{T}\,dU + \frac{p}{T}\,dV - \frac{\mu}{T}\,dn \equiv \theta\,dU + \eta\,dV - \nu\,dn\,, \tag{154}$$

defines the variables θ, η and ν and performing the Legendre transforms to one or more of these variables. One obtains in this way the so-called Massieu-Planck functions (potentials). However, since with perhaps slightly more labour, the same results can always be arrived at with the help of the potentials H, F and G, we will not discuss the Massieu-Planck functions here any more.

Returning to the usual potentials it is good to notice that although the quantities like p or μ in the differentials dU and dF are denoted by the same letter (in agreement with the usual habit of physicists to use the same symbol for the same physical quantity), they are in principle different functions of different variables (of the natural ones for the respective potentials) and their experimental determination (as mathematical functions of these variables) requires specifying the thermodynamical conditions. As an example consider a wire stretched by the force K and satisfying the Hooke's law. If its internal energy $U = U(S, L)$ is given, the tension K defined by the derivative of this function with respect to L is measured as the force needed to reversibly stretch the wire under adiathermal conditions, that is at constant entropy of the wire: $dU|_{S=\text{const}} = KdL$. It can be determined by performing a series of measurements at different values of the entropy S of the wire. The commonly used coefficient k in the Hooke's law, which in mechanics enters the formula for the wire potential energy $E_{\text{pot}} = \frac{1}{2}k(L - L_0)^2$, is however usually determined by measuring the force needed to stretch the wire at constant temperature T (equal to that of the surrounding). In such conditions the (minimal) work which must be done by a work source to stretch the wire from its equilibrium state with the length L_0 (the work which in mechanics is identified with the change of the wire's potential energy) is, as discussed in the preceding Lecture, given by ΔF at fixed T. In this way measured is the coefficient $k(T)$ in the differential (written using the Hooke's law)

$$dF = -SdT + k(T)(L - L_0)dL \equiv -SdT + K(T, L)dL\,.$$

Integrating this one-form (along the path $(T_0, L_0) \to (T, L_0) \to (T, L)$) one gets

$$F(T, L) = F(T, L_0) + \frac{1}{2}\,k(T)(L - L_0)^2\,.$$

Since $S = -(\partial F/\partial T)_L$, one obtains

$$S(T, L) = -\frac{dF(T, L_0)}{dT} - \frac{1}{2}\,\frac{dk(T)}{dT}\,(L - L_0)^2 \equiv S(T, L_0) - \frac{1}{2}\,\frac{dk(T)}{dT}\,(L - L_0)^2\,.$$

Applying now the formula (153) one obtains

$$\begin{aligned} U(T, L) &= F(T, L_0) + TS(T, L_0) + \frac{1}{2}\left(k(T) - T\,\frac{dK(T)}{dT}\right)(L - L_0)^2 \\ &= U(T, L_0) + \frac{1}{2}\left(k(T) - T\,\frac{dk(T)}{dT}\right)(L - L_0)^2\,. \end{aligned}$$

Therefore the factor k defined as the coefficient of $\frac{1}{2}(L - L_0)^2$ in the internal energy of the wire $U = U(T, L)$ (which one would naturally split into the thermal energy of the wire and its mechanical energy) is not what is measured in typical mechanical experiments conducted at constant temperature.

Consequences of extensiveness: the Gibbs-Duhem relation and chemical potentials

Reconstructing the fundamental relation of a given system that is, obtaining the full thermodynamical information about it, requires exploiting various experimental data. In the case of a simple body consisting of a fixed amount of matter one principal heat capacity at one fixed value of the body's volume or pressure must be measured as a function of temperature and the equation of state must be known (its form can be determined by measuring various differential coefficients like k_T, α_p, etc.). Together they provide sufficient information about the coefficients of the differentials in the forms $dU = C_V dT + (\partial U/\partial V)_T dV$ (or $dU = (\partial U/\partial T)_p dT + (\partial U/\partial p)_T dp$) and $dS = (C_V/T)dT + (\partial S/\partial V)_T dV$ (or $dS = (\partial S/\partial T)_p dT + (\partial S/\partial p)_T dp$) to allow to integrate them up and to obtain $U = U(T, V)$ and $S = S(T, V)$ (or $U = U(T, p)$ and $S = S(T, p)$) which is equivalent (in the case when $U = U(T, p)$ and $S = S(T, p)$ are obtained, because the equation of state has been assumed to be known) to knowing the relations $S = S(U, V)$ or $U = U(S, V)$. Reconstructing the dependence of entropy and of internal energy on each additional parameter in the case of nonsimple systems, e.g. on the magnetization M, if the work of magnetization $đW = \mathcal{H}_0 dM$ can be (reversibly) done on the system in addition to the usual volume work $-pdV$, would require determining from data one more function.[108] Similarly, reconstructing the dependence on the number(s) of moles[109] n (n_i, $i = 1, \dots, r$) would in principle require supplementary data to determine the coefficient(s) of the differential(s) dn (dn_i) in dU and/or dS. The extensiveness property of U and S (if the system can be treated as extensive) imposes, however, very stringent constraints which uniquely determine the dependence of any state function on the total amount of matter represented by the total number of moles $n = \sum_{i=1}^{r} n_i$. Thus, if an extensive system is made up of only one sort of molecules no additional data is needed beyond those related to the reversible performance of works and the reversible heat transfers. We now show this by deriving the relations which follow from extensivness of thermodynamical systems.

One of the Callenian postulates (Lecture VI) is that entropy of an extensive system is a homogeneous function of order one of its extensive arguments:

$$S(\lambda U,\, \lambda V,\, \dots, \lambda n_1, \dots,\, \lambda n_r) = \lambda\, S(U, V, \dots, n_1, \dots, n_r)\,. \tag{155}$$

The same relation applies of course to $U = U(S, V, \dots, n_1, \dots, n_r)$. One of the consequences, already discussed in Lecture VI, is that

$$S = n\, s(u, v, \dots, x_1, \dots, x_r)\,, \qquad U = n\, u(s, v, \dots, x_1, \dots, x_r)\,,$$

where u, s, etc., are molar quantities (pertaining to one mole of the substance) and $x_i = n_i/n$, where $n = \sum_{i=1}^{r} n_i$, are the molar fractions. It follows that the intensive parameters T, p, μ_i which are defined as derivatives of U with respect to the extensive

---

[108] All such functions are called equations of state. Similarly as e.g. C_V, the dependence on the volume of which in the case of simple systems is entirely determined by the equations of state, they must satisfy definite consistency conditions (following from the fact that dU and dS must be closed forms, U and S being the state functions) which, however, do not determine them completely, leaving room for experimental input. Conversely, if these functions are determined by measuring properties of real systems obeying the laws of thermodynamics, it is assumed that the obtained information is consistent, that is, the forms dU and dS reconstructed on this basis are closed.

[109] Recall that some systems not characterized by this (these) variables.

parameters must be homogeneous functions of order zero:

$$\begin{aligned} T &= T(\lambda S,\, \lambda V,\, \dots, \lambda n_1, \dots) = T(S, V, \dots, n_1, \dots) \qquad = T(s, v \dots, x_1, \dots)\,, \\ p &= p(\lambda S,\, \lambda V,\, \dots, \lambda n_1, \dots) \ = p(S, V, \dots, n_1, \dots, n_r) \ = p(s, v \dots, x_1, \dots)\,, \\ \mu_i &= \mu_i(\lambda S,\, \lambda V,\, \dots, \lambda n_1, \dots) = \mu_i(S, V, \dots, n_1, \dots, n_r) = \mu_i(s, v \dots, x_1, \dots)\,. \end{aligned} \tag{156}$$

If the number of works which can be reversibly done on the system is $o-1$, the middle equation is replaced by $o-1$ analogous equations expressing the order zero homogeneity of the corresponding generalized forces y_i. In all there are then $1+(o-1)+r$ equalities (156). Since $x_1 + \dots + x_r = 1$, the $1+(o-1)+r$ intensive parameters T, y_j, μ_i ($j = 1, \dots, o-1$, $i = 1, \dots, r$) depend on only $1+(o-1)+r-1 = (o-1)+r$ variables, which implies that there must be one relation linking the intensive variables T, $y_1, \dots, y_{o-1}, \dots, \mu_1, \dots, \mu_r$. If there is only one material component ($r = 1$), this relation uniquely determines the single chemical potential μ as a function of the $1+(o-1)$ intensive parameters T, p, ...

Furthermore, differentiating the relation (155) and the analogous relation written for U with respect to λ and setting then $\lambda = 1$, one obtains that

$$U(S, V, n_1, \dots) = ST(S, V, n_1, \dots) - V\,p(S, V, n_1, \dots) + \sum_{j=1}^{r} n_j \mu_j(S, V, n_1, \dots)\,, \tag{157}$$

$$S(U, V, n_1, \dots) = U\theta(U, V, n_1, \dots) + V\,\eta(U, V, n_1, \dots) - \sum_{j=1}^{r} n_j \nu_j(U, V, n_1, \dots)\,,$$

(recall that $\theta \equiv 1/T$, $\eta \equiv p/T$ and $\nu_j \equiv \mu_j/T$). Writing now the differential of U in its natural variables first using $U = U(S, V, n_1, \dots)$ and then using the form (157) of U we get

$$dU = TdS - p\,dV + \sum_{j=1}^{r} \mu_j dn_j\,,$$

$$dU = TdS + SdT - p\,dV - V\,dp + \sum_{j=1}^{r} \mu_j dn_j + \sum_{j=1}^{r} n_j d\mu_j\,.$$

Subtracting now these two expressions side by side leads to the formula

$$SdT - V\,dp + \sum_{j=1}^{r} n_j d\mu_j = 0\,, \tag{158}$$

known as the **Gibbs-Duhem relation**. Analogous operations done on the entropy written in two different ways give

$$Ud\Big(\frac{1}{T}\Big) + Vd\Big(\frac{p}{T}\Big) - \sum_{j=1}^{r} n_j d\Big(\frac{\mu_j}{T}\Big) = 0\,. \tag{159}$$

Dividing (158) by the total number n of moles gives the correlation of the differentials of the intensive parameters which depends only on molar quantities:

$$s\,dT - v\,dp + \sum_{j=1}^{r} x_j d\mu_j = 0\,. \tag{160}$$

This expresses the already noted fact that of $1 + (o-1) + r$ intensive parameters characterizing an extensive system, only $(o-1)+r$ are independent (therefore their differentials must be linearly dependent). If there is only one component, so that $x_1 = 1$, and the molar entropy s and molar volume v are known as functions of the temperature T and pressure p, the relation (160) written in the form

$$d\mu = -s(T,p)\,dT + v(T,p)\,dp\,, \tag{161}$$

can be, if the functions $s(T,p)$ and $v(T,p)$ are given explicitly, integrated to give the chemical potentials $\mu(T,p)$ up to a constant (the chemical potential at some reference values T_0, p_0).

From the form (157) of the internal energy and the definitions (145)-(151) it immediately follows that

$$\begin{aligned}
H(S,p,n_1,\dots) &= ST(S,p,n_1,\dots) + \sum_{j=1}^{r} n_j\mu_j(S,p,n_1,\dots)\,,\\
F(T,V,n_1,\dots) &= -Vp(T,V,n_1,\dots) + \sum_{j=1}^{r} n_j\mu_j(T,V,n_1,\dots)\,,\\
G(T,p,n_1,\dots) &= \sum_{j=1}^{r} n_j\mu_j(T,p,n_1,\dots)\,, \qquad (162)\\
\Omega(T,V,\mu_1,\dots) &= -Vp(T,V,\mu_1,\dots)\,,
\end{aligned}$$

The same relations follow, of course, from the scaling properties of these functions: e.g. differentiating with respect to λ the relation

$$F(T,\lambda V,\,\lambda n_1,\dots) = \lambda F(T,V,n_1,\dots)\,,$$

and setting $\lambda = 1$ the second relation (162) is obtained. If there is only one component in the system, the third of these relations, after division by n gives

$$\mu(T,p) = g(T,p) \equiv \frac{1}{n}\,G(T,p,n) = u - T\,s + p\,v\,. \tag{163}$$

As advertised in the preceding Lecture, in this case the chemical potential is just the molar Gibbs function. The formula (163) gives another method of calculating the chemical potential of a system composed of one component only.

Two important facts concerning systems composed of more than one material component ($r > 1$) deserve to be clearly stated here. If such a system is closed (does not exchange matter with its surrounding), its composition, that is the numbers $n_1, \ldots, n_r$ of moles of its material constituents (chemical compounds), can change only due to chemical reactions occuring in it when the other parameters (U, $V, \ldots$ when the system is only in mechanical contact with its surrounding, or T, $V, \ldots$, when it is in thermal contact with a heat bath at the temperature T, etc.) are varied. The constitutive variables $n_1, \ldots, n_r$ are in such a situation not entirely independent variables: once they are fixed at some values of the remaining parameters (in one equilibrium state), their values at different equilibrium states are completely determined by the values of the other parameters. Therefore, although the changes of the thermodynamic potentials between two infinitesimally close equilibrium states can formally be written in the form (taking U as an example)

$$dU = TdS - pdV + \sum_{j=1}^{r} \mu_j dn_j \,,$$

this must be equivalent to $dU = TdS - p\,dV$, because as far as its interactions with the rest of the world are concerned, a closed system can be treated as a black box. It follows that changes of the constitutive variables $n_1, \ldots, n_r$ due to chemical reactions between any two infinitesimally close equilibrium states must be such that

$$\sum_{j=1}^{r} \mu_j dn_j = 0 \,. \tag{164}$$

Of course, if the system is allowed to exchange matter with its surrounding when passing from one equilibrium state to another one, then $dn_i = dn_i^{\rm chem} + dn_i^{\rm inflow}$ and only the differentials $dn_i^{\rm chem}$ satisfy (164).

Furthermore, differentiating the form (162) of the Gibbs function with respect to the variable n_j one obtains ($n_{j'}$ stands for all the remainig constitutive variables)

$$\left(\frac{\partial G}{\partial n_j}\right)_{T,p,n_{j'}} = \mu_j + \sum_{i=1}^{r} n_i \left(\frac{\partial \mu_i}{\partial n_j}\right)_{T,p,n_{j'}} .$$

Since on the other hand $(\partial G/\partial n_j)_{T,p,n_{j'}} = \mu_j$, it follows that[110]

$$\sum_{i=1}^{r} n_i \left(\frac{\partial \mu_i}{\partial n_j}\right)_{T,p,n_{j'}} = 0 \,. \tag{165}$$

Finally, dividing the Gibbs function G of such a system written as in (162) as the sum of chemical potentials weighted by the respective mole numbers by the total number of moles $n = \sum_{j=1}^{r} n_j$ one obtains

$$g(T, p, x_1, \ldots, x_r) = \sum_{j=1}^{r} x_j \mu_j(T, p, x_1, \ldots, x_r) \,. \tag{166}$$

---

[110] This expresses mathematically the fact that the chemical potentials μ_i depend on only $r-1$ molar fractions $x_1, \ldots, x_{r-1}$, because $x_r = 1 - x_1 - \ldots - x_{r-1}$.

Of course, determination of the individual chemical potentials insuch a case requires more experimental input.

Mixture of perfect gases and the entropy of mixing
The simplest multicomponent system which can be analyzed is the mixture of r different perfect gases. Its thermodynamical functions can be explicitly written down by appealing to the so called Gibbs postulate (which replaces the experimental input) which says that the internal energy U and entropy S of such a mixture in equilibrium is, when **expressed through the temperature and the volume** in which it is enclosed, simply the sum of internal energies and entropies of the individual gases treated as independent (on account of the fact that they are mutually noninteracting) and enclosed in the same volume (we remember, however, that the internal energy of a perfect gas depends only on its temperature, and not on the volume it occupies):

$$\begin{aligned} U(T, n_1, \ldots, n_r) &= \sum_{i=1}^{r} U_i(T, n_i) = \sum_{i=1}^{r} n_i u_i(T) = n \sum_{i=1}^{r} x_i u_i(T)\,, \\ S(T, V, n_1, \ldots, n_r) &= \sum_{i=1}^{r} S_i(T, V, n_i) = \sum_{i=1}^{r} n_i s_i(T, v_i) = n \sum_{i=1}^{r} x_i s_i(T, v_i)\,, \end{aligned} \tag{167}$$

where $n = n_1 + \ldots + n_r$, $x_i = n_i/n$ and $v_i = V/n_i$.

The Gibbs postulate is consistent with the fact that in general the formulae for $U(T, V, n_1, \ldots, n_r)$ and $S(T, V, n_1, \ldots, n_r)$ must (the Callenian approach!) follow from the fundamental relation (97): since the Gibbs postulate applies to perfect gases only, it is in principle possible to invert the formula

$$U = \sum_i n_i u_i(T_0) + \sum_i n_i \int_{T_0}^{T} dT'\, c_v^{(i)}(T')\,,$$

with respect to T and to obtain in this way $S = S(U, V, n_1, \ldots, n_r)$. Moreover, the general relation $dS(U, V) = (dU + p\, dV)/T$ is also satisfied:

$$\begin{aligned} dS &= \sum_i \frac{\partial S_i}{\partial U_i}\, dU_i + \sum_i \frac{\partial S_i}{\partial V}\, dV = \sum_i \frac{1}{T_i}\, dU_i + \sum_i \frac{p_i}{T_i}\, dV \\ &= \frac{1}{T}\, d\left(\sum_i U_i\right) + \frac{1}{T}\left(\sum_i p_i\right) dV = \frac{1}{T}\, dU + \frac{p}{T}\, dV\,. \end{aligned} \tag{168}$$

The immediate consequences of the postulate are the formulae for the molar heat capacity of the mixture is the weighted sum:

$$c_v^{\rm mix} \equiv \frac{1}{n}\left(\frac{\partial U}{\partial T}\right)_V = \frac{1}{n}\sum_{i=1}^{r}\left(\frac{\partial U_i}{\partial T}\right)_V = \frac{1}{n}\sum_{i=1}^{r} n_i \left(\frac{\partial u_i}{\partial T}\right)_V = \sum_{i=1}^{r} x_i\, c_v^{(i)}\,,$$

and its pressure (we remember the formula for $S(T, V, n)$ of the perfect gas!) is given by

$$\frac{p^{\rm mix}}{T} \equiv \left(\frac{\partial S}{\partial V}\right)_{U, n_i} = \sum_{i=1}^{r}\left(\frac{\partial S_i}{\partial V}\right)_{U, n_i} = \sum_{i=1}^{r} n_i \frac{R}{V} = \frac{1}{T}\sum_{i=1}^{r} p_i\,,$$

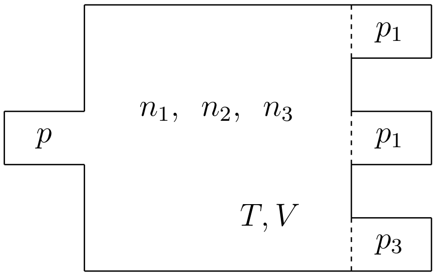

Figure 19: Comparison of the total and partial pressures of a mixture of three gases using the van 't Hoff box. Dotted lines represents membranes permeable to molecules of only one gas each.

because the derivative at constant internal energy U is here (that is for perfect gases!) the same as the derivative at constant temperature T. As a consequence of the Gibbs postulate, the pressure of a mixture of perfect gases is therefore, consistently with (168) and in agreement with the known Dalton law, the sum of the so-called partial pressures, that is the pressures which each of the gases would individually exert on the walls of the container in the absence of the other constituent gases (we drop from now the subscript "mix"):

$$p = \sum_{i=1}^{r} p_i = \sum_{i=1}^{r} n_i \frac{RT}{V} \equiv n \frac{RT}{V}\,.$$

Both these consequences of the Gibbs postulate can be directly verified: the first one by trivial measurement of the molar heat capacities of individual gases and of their mixture and the second one by the device called the van 't Hoff box shown in Figure 19 - which in principle allows to measure and compare the total and partial pressures. The Gibbs postulate finds its justification in statistical mechanics - the case of a mixture of perfect gases, made of molecules the mutual interactions of which are negligibly weak[111] is solvable within the so called canonical ensemble approach which gives the Helmholtz function $F(T, V, n_1, \dots, n_r)$ of the system (which, as we already know, contains complete information about its termodynamical characteristics): in the case of nointeracting gases it automatically gives F of the form

$$F(T, V, n_1, \dots, n_r) = \sum_{i=1}^{r} F_i(T, V, n_i)\,,$$

from which the Gibbs postulate readily follows.

---

[111] But some interactions must be present because otherwise the mixture of gases could never come to equilibrium and different gases in the mixture could have different temperatures as it happens in the present day Universe in which perfect and mutually noninteracting gases of photons and of neutrinos do indeed have different temperatures.

One more consequence of the Gibbsa postulate is the **entropy of mixing**, which is easily identified when the formula (167):

$$S(T,V,n_1,\dots,n_r)=\sum_{i=1}^{r}n_i\left(s_i(T_0,v_0)+\int_{T_0}^{T}\frac{dT'}{T'}\,c_v^{(i)}(T')+R\ln\frac{v_i}{v_0}\right),$$

for the entropy of the mixture (as follows from the perfect gas equation of state, molar volumes v_0 of all perfect gases taken at the same reference temperature and pressure are equal) is explicitly expressed (using the relations $v_i=V/n_i$) through the total volume V occupied by the mixture and the total mole number n:

$$\begin{aligned}S(T,V,n,x_1\dots,x_r)=n\,s^{\text{mix}}(T_0,v_0)+n\int_{T_0}^{T}\frac{dT'}{T'}\,c_v^{\text{mix}}(T')+nR\ln\frac{V}{nv_0}\\ +n\sum_{i=1}^{r}(-R\,x_i\ln x_i)\,.\end{aligned}\tag{169}$$

(We have introduced here $s^{\text{mix}}(T_0,v_0)=\sum_i x_is_i(T_0,v_0)$.) The last term is the mixing entropy which we will also call the Cinderella's (Kopciuszkowa) entropy.[112] It is positive because $x_i\le 1$. That this term is related to the mixing of different gases in one container follows clearly from the comparison of (169) with the entropy of n moles of a one-component perfect gas which (accidentally) would have the same molar heat capacity and the same molar entropy s_0 as the mixture:

$$S(T,V,n)=ns(T_0,v_0)+n\int_{T_0}^{T}\frac{dT'}{T'}\,c_v(T')+nR\ln\frac{V}{nv_0}\,.$$

The mixing entropy can be seen in many ways (hopefully they will be discussed in classes). The one which offers an insight into its origin is as follows. Prepare n_1 and n_2 moles of two different (perfect) gases in two initially isolated containers at the same pressure and the same temperature. If the two containers are joined together, each of the gases expands freely into the additional volume (undergoing essentially the Joule process). The resulting entropy change is just the mixing entropy. It should be however noticed, that if the two gases were identical, one would say that being at the same temperature and pressure they are in equilibrium and when the two containers are joined, nothing happens - the total entropy does not increase. This shows that the mixing entropy results from different treatment of same and different gases and this procedure finds its justification only in statistical physics and in fact in the quantum mechanical indistinguishability of identical particles.

Since the Gibbs postulate allows to explicitly construct all thermodynamical functions of the mixture of perfect gases, also the chemical potentials can be obtained. It is a matter of a simple calculation to find that

$$\mu_i(T,p,x_1,\dots,x_r)=\mu(T,\,x_ip)\equiv\mu(T,p_i)=\mu(T,p)+RT\ln x_i\,,\tag{170}$$

---

[112] Everyone knows that the task assigned to Cinderella by her bad step mother was to separate poppy from pea, that is precisely to reduce the mixing enetropy.

where $\mu(T,p)$ is the chemical potential of a one-component perfect gas at temperature T and pressure p.

Another situation in which some approximate formulae for the chemical potentials of two material components of the system can be obtained without additional experimental (or statistical physics) input is one in which the number of moles of one of the two components is much smaller than that of the other one. Let $x = n_2/(n_1+n_2) \ll n_1/(n_1+n_2) = 1-x$. The molar internal energy and the molar volume of such a weak solution, which can be treated as functions od T, p and x, can be then expanded in the Taylor series in $x \ll 1$:

$$\begin{aligned}u(T,p,x) &= u_0(T,p) + x\,\Delta u(T,p) + \dots\,,\\ v(T,p,x) &= v_0(T,p) + x\,\Delta v(T,p) + \dots\,,\end{aligned}$$

in which u_0 and v_0 are the molar internal energy and volume of the pure solvent (the other component is called the solute). The differential of the molar entropy (taken at constant x) can be then organized as follows

$$ds(T,p,x) = \left\{\frac{1}{T}\,du_0(T,p) + \frac{p}{T}\,dv_0(T,p)\right\} + x\left\{\frac{1}{T}\,d\Delta u(T,p) + \frac{p}{T}\,d\Delta v_0(T,p)\right\}.$$

As a whole, the right hand side must be an exact differential (as the differential of the molar entropy of the complete solution) in the variables T and p. Moreover, the first bracket must also be an exact differential of the molar entropy of the pure solvent. It follows, that the second bracket must too have this property and the one-form ds can be integrated up in the variables T and p yielding

$$s(T,p,x) = s_0(T,p) + x\,\Delta s(T,p) + f(x)\,,$$

where $f(x)$ is an integration constant (which must, therefore, be independent of T and p). One then invokes the Planck argument that at sufficiently high temperatures all substances turn into gases, and ultimately (at low pressures) behave as perfect gases. This means that the integration constant $f(x)$ must be (as the only term which is independent of T and p) the mixing entropy of the mixture of two perfect gases:

$$f(x) = -R(1-x)\ln(1-x) - Rx\ln x\,.$$

This allows to construct the molar Gibbs function of the solution:

$$g(T,p,x) = u - Ts + p\,v = g_0(T,p) + x\Delta g(T,p) + RT\left\{(1-x)\ln(1-x) + x\ln x\right\},$$

in which $g_0 = u_0 - Ts_0 + pv_0$ and $\Delta g_0 = \Delta u_0 - T\Delta s_0 + p\Delta v_0$. The extensive Gibbs function of $n = n_1 + n_2$ moles of the solution is now constructed using the standard prescription ($nx = n_2$, $n(1-x) = n_1$)

$$\begin{aligned}G(T,p,n_1,n_2) &= (n_1+n_2)\,g(T,p,x)\\ &= (n_1+n_2)\,g_0(T,p) + n_2\Delta g(T,p) + RTn_1\ln\frac{n_1}{n_1+n_2} + RTn_2\ln\frac{n_2}{n_1+n_2}\,.\end{aligned}$$

The chemical potentials can be then obtained as the derivatives

$$\mu_1 = \left(\frac{\partial G}{\partial n_1}\right)_{T,p,n_2} = g_0(T,p) + RT\ln(1-x) \approx g_0(T,p) - RT\,x\,,$$
$$\mu_2 = \left(\frac{\partial G}{\partial n_2}\right)_{T,p,n_1} = g_0(T,p) + \Delta g(T,p) + RT\ln x \equiv \psi(T,p) + RT\ln x\,. \qquad (171)$$

While the temperature dependence of μ_2 (the chemical potential of the solute) is usually unknown (the one of μ_1 is the same as of the pure solvent), its dependence on the concentration $x \ll 1$ is obtained explicitly.

LECTURE IX (TMD)

Equilibrium of phases

By phases of a substance one understands different (spatially distinguished) forms in which it may exist. Different phases of the same substance have different physical properties. (It will become clear that the first part of this tentative definition not always applies and it is rather the second part which is more appropriate.) In some range of intensive parameters characterizing a given substance a given its phase is stable in another this phase may become metastable and in yet some other ranges it simply cannot exist. If in certain conditions usually characterized by the values of temperature and pressure (and of other intensive parameters, if the substance is not simple) more than one phase may exist as stable, one speaks of the coexistence of phases. The system is then inhomogeneous - the phases exist as spatially separated forms. (Again, this, as it will become clear, applies only to one class of phase transitions.) They are then treated as different parts of a compound system separated by (fictitious) walls allowing for a transfer of matter, energy and volume between the parts.[113] It is in these conditions that phase transitions normally occur.[114]

Phases of a given substance should not be confused with its physical states (solid, liquid and gaseous). The solid state of a given substance can have different phases: for example, silicon (Si) has 3 different phases (called allotropic modifications) in which atoms are differently ordered in the crystalline lattice cells; tin (Sn) can exist in two phases (white tetraghonal - a crystalline one - and grey which is amorphous) both of which can be treated as a solid state; H_2O has at least 12 different crystalline modifications each of which is a separate ice phase! There are also known different mixtures of liquids (different phases of the mixture): one phase rich in one component and another phase rich in another component.

The conditions in which two different phases of a simple one-component substance can coexist were already discussed at the end of Lecture VII: independently of whether the system was isolated or open (interacting in this or another way with its surrounding), the coexistence of two phases α and β of such a substance always required the equality of the temperatures (thermal equilibrium with respect to the exchange of heat), pressures (mechanical equilibrium with respect to the exchange of volume) and the equality of the chemical potentials (of the single matter component) of the coexisting phases.[115] The

---

[113] The boundary separating the phases, although neglected in the further discussion here, has in fact some thickness; it can be treated as consisting of some quantity (negligible, from the thermodnamic point of view, compared to the amount of matter in the phases themselves) of matter, and ascribed entropy, energy and other thermodynamic functions. In this way it can be included into the analysis as yet another part of the system. Its properties (e.g. the surface tension) can also lead to modifications of the equilibrium condition; these can also be modified in presence of external fields like the gravitational one.

[114] A phase which in a given range of the parameters is metastable may exist for a very long time but eventually turns out into the stable (in this range) phase. This means that the change of one phase into another one can also occur in conditions in which only a single phase is stable.

[115] However, the character of the equilibrium - absolutely stable or neutral - does depend on whether the system is isolated or is under prescribed pressure and temperature.

considerations of Lecture VII can be easily generalized, assuming for convenience that the system is isolated, to f phases labeled by the index $\alpha = 1, \dots, f$ and r material components (labeled by the index i). If a closed system composed of several phases is isolated from its surrounding, it is the maximum entropy principle which determines its equilibrium state. Therefore, the variation

$$\begin{aligned}\delta S &= \sum_{\alpha=1}^{f} \delta S^{(\alpha)}(U^{(\alpha)}, V^{(\alpha)}, n_1^{(\alpha)}, \dots, n_r^{(\alpha)}) \\ &= \sum_{\alpha=1}^{f} \left\{ \frac{\partial S^{(\alpha)}}{\partial U^{(\alpha)}}\, \delta U^{(\alpha)} + \frac{\partial S^{(\alpha)}}{\partial V^{(\alpha)}}\, \delta V^{(\alpha)} + \sum_{i=1}^{r} \frac{\partial S^{(\alpha)}}{\partial n_i^{(\alpha)}}\, \delta n_i^{(\alpha)} \right\},\end{aligned}$$

representing departures of the system's entropy (in virtual equilibrium states which can be realized with the help of suitable walls, i.e. of stronger internal constraints, blocking the matter, energy and volume transfers between the phases) from its true equilibrium value, subjected to the restrictions (we assume for the moment that chemical reactions are not allowed to occur)

$$\sum_{\alpha=1}^{f} \delta U^{(\alpha)} = 0\,, \qquad \sum_{\alpha=1}^{f} \delta V^{(\alpha)} = 0\,, \qquad \sum_{\alpha=1}^{f} \delta n_i^{(\alpha)} = 0\,, \;\; i = 1, \dots, r\,, \tag{172}$$

must vanish. Introducing as previously the Lagrange multipliers written as $1/T$, p/T and $-\mu_i/T$ allows (see Lecture IV, p. 61 for explanation) to treat in the condition

$$\sum_{\alpha=1}^{f} \left\{ \left[\frac{\partial S^{(\alpha)}}{\partial U^{(\alpha)}} - \frac{1}{T}\right] \delta U^{(\alpha)} + \left[\frac{\partial S^{(\alpha)}}{\partial V^{(\alpha)}} - \frac{p}{T}\right] \delta V^{(\alpha)} - \sum_{i=1}^{r} \left[-\frac{\partial S^{(\alpha)}}{\partial n_i^{(\alpha)}} - \frac{\mu_i}{T}\right] \delta n_i^{(\alpha)} \right\} = 0\,,$$

all variations $\delta U^{(\alpha)}$, $\delta V^{(\alpha)}$ and $\delta n_i^{(\alpha)}$ as effectively independent. This leads to the conditions[116]

$$\frac{\partial S^{(\alpha)}}{\partial U^{(\alpha)}} = \frac{1}{T}\,, \qquad \frac{\partial S^{(\alpha)}}{\partial V^{(\alpha)}} = \frac{p}{T}\,, \qquad -\frac{\partial S^{(\alpha)}}{\partial n_i^{(\alpha)}} = \frac{\mu_i}{T}\,, \qquad i = 1, \dots, r\,, \tag{173}$$

which must be satisfied by all the f phases. Since in effect the multipliers $1/T$, p/T and $-\mu_i/T$ correlate the departures $\delta S^{(\alpha)}$, $\delta U^{(\alpha)}$, $\delta V^{(\alpha)}$ and $\delta n_i^{(\alpha)}$ of entropies, energies, volumes and numbers of moles (from the values of these parameters in the true equilibrium state) of possible virtual equilibrium states of the phases so that

$$\delta S^{(\alpha)} = \frac{1}{T}\, \delta U^{(\alpha)} + \frac{p}{T}\, \delta V^{(\alpha)} - \sum_{i=1}^{r} \frac{\mu_i}{T}\, \delta n_i^{(\alpha)}\,, \tag{174}$$

---

[116] Of course, to arrive at the equilibrium conditions (173) one could use as well the principle of minimum internal energy (at constant entropy).

they acquire the interpretations of the equilibrium inverse temperature, pressure and chemical potentials (divided by the temperature) which must be common for all the phases. In all, there are, therefore, $(2+r)f$ relations (173) which together with $2+r$ constancy conditions (172) provide the equations allowing to completely determine $(2+r)f$ equilibrium values of the variables $U^{(\alpha)}$, $V^{(\alpha)}$, $n_1^{(\alpha)}, \dots, n_r^{(\alpha)}$, $\alpha = 1, \dots, f$ and $2+r$ Lagrange multipliers. In principle the first two sets of the equations (173) allow to parametrize the internal energies $U^{(\alpha)}$ and volumes $V^{(\alpha)}$ of the phases in terms of the common temperature and pressure T, p and the numbers of moles $n_1^{(\alpha)}, \dots, n_r^{(\alpha)}$ in these phases; the last set of conditions (173), the equalities between the phases of the corresponding chemical potentials, should be then solved for T, p and $n_1^{(\alpha)}, \dots, n_r^{(\alpha)}$ respecting the constancy of the total energy U, volume V and the total numbers n_i of the matter components in the entire isolated system.[117] We have assumed here that all coexisting phases may exist with arbitrary concentrations of their r matter components and that all these matter components may flow from one phase to another one. If this is not the case, (for instant some matter components cannot be present in some phases) the conditions must be rederived taking the restrictions into account.

Treating $U^{(\alpha)}$ and $V^{(\alpha)}$ as functions of the (common) temperature and pressure one can also relax the conditions of constancy of U and V and operate with T and p as the fundamental variables. This is equivalent to considering a closed system which is not isolated but remains in thermal and mechanical contact with the surrounding at temperature T and pressure p; in this case one could also appeal to the general stability requirement (minimum of the system's availability A discussed in Lecture VII) which dictates that in equilibrium the system's temperature (common to all phases) and pressure (also common to all phases) must be equal to those of the surrounding and then, at fixed T and p, seek the minimum of the system's Gibbs function

$$\begin{aligned} G(T,p,n_1^{(1)},\dots,n_r^{(f)}) &= \sum_{\alpha=1}^{f} G^{(\alpha)}(T,p,n_1^{(\alpha)},\dots,n_r^{(\alpha)}) \qquad (175) \\ &\equiv \sum_{\alpha=1}^{f}\sum_{i=1}^{r} n_i^{(\alpha)}\mu_i^{(\alpha)}(T,p,x_1^{(\alpha)},\dots,x_r^{(\alpha)})\,, \end{aligned}$$

with respect to the variations $\delta n_i^{(\alpha)}$ subjected to r constancy conditions. Introducing again r Lagrange multpliers which acquire the interpretation of the common for all phases r chemical potentials μ_i, $i = 1, \dots, r$, one readily finds (using the relation (165) which must be satisfied for each of the phases) the equilibrium conditions in the form $\mu_i^{(\alpha)}(T,p,x_1^{(\alpha)},\dots,x_r^{(\alpha)}) = \mu_i$, which at given T and p must be solved together with the r conditions of constancy of the total numbers of moles $n_i = n_i^{(1)} + \dots + n_i^{(f)}$ of r constituents. The difference with the case of the completely isolated system is that now the total system's volume V and its total energy U are not fixed and the equilibrium conditions may not determine the distribution of all the r numbers of moles n_i between the

---

[117] Obviously, carrying out this programme in practice requires that the fundamental relations of all the f phases, i.e. the functional dependence of the entropies $S^{(\alpha)}$ on their natural arguments be known.

coexisting phases (and the volumes occupied by these phases) - the equilibrium may be of the neutral character.

It is worth to generalize the above considerations by allowing chemical reactions to occur in a system closed as a whole and composed of f phases and r material components (different chemical compounds A_i, $i = 1, \dots, r$). Again to establish the necessary conditions for the equilibrium we will treat the system as isolated. Of course, if one restricts to a single phase ($f = 1$), the conditions obtained below will apply to chemical reactions taking place in a single homogeneous system. A given allowed reaction written in the form (here we assume that A_i include all possible chemical compounds: those which are present in the initial state and those which form the final state)

$$\sum_{i=1}^{r} \nu_i A_i \longrightarrow \sum_{j=1}^{r} \nu'_j A_j \,, \tag{176}$$

where ν_i and ν'_j are the (in this notation positive) stoichiometric coefficients (numbers of moles of the compounds involved in a single act of the reaction) will be now written in the more convenient form

$$\sum_{j=1}^{r} \nu_j A_j = 0 \,, \tag{177}$$

in which the stoichiometric coefficients of the products are (by convention) counted as positive and those of the reagents as negative. With this convention a single act of the reaction changes the number n_i of moles of the i-th chemical compound in the system by ν_i. It follows that the changes δn_i of the total number of moles (in all phases - since all the chemical compounds are allowed here to flow from one phases to another one, the phase in which the reactions actually take place has no significance) of the i-th chemical compound is given by $\delta n_i = \nu_i \, \delta\zeta$, where $\delta\zeta$, is the factor correlating the possible changes of the number of moles of different compounds.

If more independent chemical reactions

$$\sum_{i=1}^{r} \nu_i^{(a)} A_i = 0 \,, \qquad a = 1, \dots, R \,, \tag{178}$$

(with $\nu_i^{(a)}$ the stoichiometric coefficients of the a-th reaction - some of them can be zero) between the compounds of the cosidered system are possible,[118] the changes δn_i of the total numbers of moles are given by

$$\delta n_i \equiv \sum_{\alpha} \delta n^{(\alpha)} = \sum_{a=1}^{R} \nu_i^a \delta\zeta^a \,. \tag{179}$$

---

[118] The reactions are independent if the R equations (178) treated as linear homogeneous equations for the A_i's are linearly independent. This means that their number R cannot exceed the number r of the chemical compounds involved (otherwise the rank of the matrix formed by the stoichiometric coeffcients would certainly be less than R).

Allowing for chemical reactions in seeking the equilibrium state of an isolated closed system which may exist in f phases amounts, therefore, to replacing the r conditions $\sum_\alpha \delta n_i^{(\alpha)} = 0$ by the r conditions

$$\sum_\alpha \delta n_i^{(\alpha)} - \sum_{a=1}^{R} \nu_i^{(a)}\, \delta\zeta^{(a)} = 0\,, \qquad i = 1, \dots, r\,. \tag{180}$$

Adding them together with the conditions of constancy of U and V with the Lagrange multipliers $-\mu_i/T$, $1/T$ and p/T to the variation of the entropy gives as the equilibrium condition the equality

$$\sum_{\alpha=1}^{f} \left\{ \left[\frac{\partial S^{(\alpha)}}{\partial U^{(\alpha)}} - \frac{1}{T}\right] \delta U^{(\alpha)} + \left[\frac{\partial S^{(\alpha)}}{\partial V^{(\alpha)}} - \frac{p}{T}\right] \delta V^{(\alpha)} \right\}$$
$$- \sum_{i=1}^{r} \sum_{\alpha=1}^{f} \left(-\frac{\partial S^{(\alpha)}}{\partial n_i^{(\alpha)}}\, \delta n_i^{(\alpha)} \right) + \sum_{i=1}^{r} \frac{\mu_i}{T} \left(\sum_{\alpha=1}^{f} \delta n_i^{(\alpha)} - \sum_{a=1}^{R} \nu_i^{(a)}\, \delta\zeta^{(a)} \right) = 0\,.$$

This upon a rearrangement and taking into account that owing to the Lagrange multipliers all variations $\delta U^{(\alpha)}$, $\delta V^{(\alpha)}$, $\delta n_i^{(\alpha)}$ and $\delta\zeta^a$ can effectively be treated as independent, leads as previously to the equality of the temperatures and pressures of all the phases, to the equality of the chemical potentials of each of the matter components in all the phases: $\mu_i^{(1)} = \ldots = \mu_i^{(f)} \equiv \mu_i(T, p)$ - all the functions $\mu_i^{(\alpha)}(T, p, x_1^{(\alpha)}, \dots, x_r^{(\alpha)}) = (\partial S^{(\alpha)}/\partial n_i^{(\alpha)})$, $\alpha = 1, \dots, f$ assume a common value[119] $\mu_i(T, p)$ - and to the R additional conditions[120]

$$\sum_{i=1}^{r} \nu_i^{(a)}\, \mu_i(T, p) = 0\,, \qquad a = 1, \dots, R, \tag{181}$$

correlating the values $\mu_i(T, p)$ of the chemical components taking part in the allowed reactions. In all, there are in this case $(2 + r)f$ relations (173) which together with $2 + r$ conditions $\sum_\alpha \delta U^{(\alpha)} = 0$, $\sum_\alpha \delta V^{(\alpha)} = 0$ and (180) as well as R conditions (181) provide just the right number of equations alowing to completely determine $(2 + r)f$ equilibrium values of the variables $U^{(\alpha)}$, $V^{(\alpha)}$, $n_1^{(\alpha)}, \dots, n_r^{(\alpha)}$, $\alpha = 1, \dots, f$ and $2 + r + R$ Lagrange multipliers T, p, μ_i and ζ^a.

---

[119] Again, since T, p and μ_i correlate as in (174) the infinitesimal differences $\delta U^{(\alpha)}$, $\delta V^{(\alpha)}$ and $\delta n_i^{(\alpha)}$ of the parameters of the phase α between its real and virtual equilibrium states, they acquire the interpretation of the values of the system's common temperature, pressure and chemical potentials.

[120] It is precisely the relation (181), which in equilibrium must be satisfied by the chemical potentials of every phase, combined with the relations $dn_i = \sum_a \nu_i^{(a)} d\zeta^{(a)}$ which implies that when a phase, treated now as a separate system, is in equilibrium in itself and then passes to another neighbouring equilibrium state as a result of an infinitesimal change of the conditions (e.g. a change of the external pressure and temperature, if it is treated as being in thermal and mechanical contact with its surrounding), the changes dn_i of its material constituents induced by possible chemical reactions are such that $\sum_i \mu_i dn_i = 0$ - there is no change of the system's Gibbs function due to the reaction. This, as already explained in Lecture VIII, is consistent with the possibility, fundamental for classical thermodynamics, of treating the system as a black box, i.e. of treating its thermodynamic potentials as functions of external variables only.

An analogous reasoning can be also carried out (with the same conclusion as to the equilibrium conditions, though perhaps not as far as the type of stability of the equilibrium state is concerned) treating the closed system consisting of f phases as being in thermal and mechanical contact with its surrounding at the temperature T and pressure p and minimizing the system's Gibbs function (175) with respect to $n_i^{(\alpha)}$ respecting the conditions (180).

Gibbs phase rule

If no chemical reactions are allowed, the equilibrium condition found in Lecture VII extends straightforwardly to multicomponent and multiphase systems, the only change being that in such cases the chemical potentials characterizing the same component in any of the phases must be equal:

$$\mu_i^{(\alpha)}(T, p, x_1^{(\alpha)}, \dots, x_{r-1}^{(\alpha)}) = \mu_i^{(\beta)}(T, p, x_1^{(\beta)}, \dots, x_{r-1}^{(\beta)}), \qquad i = 1, \dots, r. \tag{182}$$

The question (left out in the foregoing considerations) how many phases of a given substance can simultaneously be in equilibrium, that is when the derived conditions can be satisfied and if they can, how much freedom there is in satisfying them, finds the answer in the **Gibbs phase rule** which we derive now. If each phase is treated as a separate closed (sub)system, it is characterized by $2 + r$ extensive parameters: its energy $U^{(\alpha)}$, its volume $V^{(\alpha)}$ and the numbers of moles $n_1^{(\alpha)}, \dots, n_r^{(\alpha)}$ of the components. The first two variables can be traded for the temperature $T^{(\alpha)}$ and pressure $p^{(\alpha)}$ of each phase. In all, f isolated phases are characterized therefore by $(2 + r) \cdot f$ independent parameters. If the phases are to be in equilibrium, their intensive parameters must satisfy a number of conditions. As has been established above, they must, firstly, have all the same temperatures (thermal equilibrium)

$$T^{(1)} = T^{(2)} = \dots = T^{(f)} \equiv T,$$

which gives $f - 1$ conditions. Another $f - 1$ conditions follow from the equality of their pressures (mechanical equilibrium):

$$p^{(1)} = p^{(2)} = \dots = p^{(f)} \equiv p.$$

Equalities of the chemical potentials (chemical equilibrium with respect to matter exchange)

$$\begin{gathered} \mu_1^{(1)} = \mu_1^{(2)} = \dots = \mu_1^{(f)} \equiv \mu_1, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \mu_r^{(1)} = \mu_r^{(2)} = \dots = \mu_r^{(f)} \equiv \mu_r, \end{gathered}$$

give together $(f - 1) \cdot r$ further conditions. Finally, there are f conditions

$$\sum_{i=1}^{r} x_i^{(1)} = 1, \qquad \dots, \qquad \sum_{i=1}^{r} x_i^{(f)} = 1,$$

following from the fact that the equilibrium of phases depends only on the values of the intensive parameters which in turn do not depend on all extensive parameters but only on their ratios (see Lecture VIII). This is taken into account by expressing them in terms of the molar fractions x_i which in each phase sum up to unity. Therefore there remains the freedom of varying

$$k = (2+r)\cdot f - [(2+r)\cdot(f-1) + f] = 2 + r - f\,, \tag{183}$$

independent parameters.[121] Of course, the coexistence of phases is possible only if

$$k \equiv 2 + r - f \geq 0\,.$$

Thus two phases ($f = 2$) of a single component ($r = 1$) substance (like pure H_2O) can coexist along a curve because $k = 2 + 1 - 2 = 1$ which means that one parameter can be varied freely, while three phases ($f = 3$) of a single-component substance can coexist only at an isolated point (isolated points). If there are more components, more phases can coexist at one point, more along a curve etc. Finally, if chemical reactions are allowed to occur between the components of the system, each possible reaction imposes one more condition (181) on the chemical potentials that is, on the molar fractions on which they depend, and the number of degrees of freedom decreases by one per each reaction.[122]

Chemical reactions in a single phase.
The general conditions of equilibrium of arbitrary multiphase systems composed of many different constituents (chemical compounds) A_i, $i = 1, \dots, r$ between which R independent chemical reactions (178) can occur, established in this lecture can be straightforwardly adapted to the case of a closed single phase system and read:

$$\sum_{i=1}^{r} \nu_i^{(a)} \mu_i = 0\,, \qquad a = 1, \dots, R\,. \tag{184}$$

---

[121] It should be noticed that in this reasoning we allow for a freedom of adjusting the system's (composed of f phases) total energy and volume (or equivalently - if it is treated as being in thermal and mechanical contact with its surrounding - T and p) as well as the total numbers n_i of its constituents. We are now not asking wheter in given conditions (set by given values of U, V and n_i in the case of an isolated system, or by T, p and n_i in the case of a closed system in thermal and mechanical equilibrium with its surrounding) the equilibrium of phases will occur, but rather what is the freedom in adjusting these global (external) parameters to obtain the equilibrium. Another way of arriving at the result (183) is as follows: the intensive parameters which are at play are: T and p which must be common to all phases and $f\cdot(r-1)$ independent molar fractions $x_i^{(\alpha)}$ in all f phases; the equalities of the chemical potentials constitute $r\cdot(f-1)$ conditions which must be satisfied if f phases coexist in equilibrium. This again gives $2 + r - f$ free parameters.

[122] Once again it should be stressed that the problem here is how much freedom in satisfying the requirements of equilibrium there is in general; for this reason arbitrary (equilibrium) concentrations of components in different phases are admited here and the fact that in a (multiphase) system closed as a whole changes of the total numbers of moles of the constituents due to the chemical reactions are constrained by the conditions (180) is not relevant. The question what are the final (equilibrium) numbers $n_i^{(\alpha)}$ of moles of different constituents in different phases if one starts with fixed initial total numbers $n_i^{\rm init}$ of moles of the constituents in certain definite conditions (fixed U and V or fixed T and p) and let them to react is a different one. (These different questions seems to be confused in the presentation of Werle).

These conditions must be satisfied irrespectively of the conditions (absolutely isolated system - with its total energy U and total volume V fixed, a closed system in thermal contact with its surrounding - fixed T and V, or in thermal and mechanical contact in which case fixed are its temperature T and pressure p) in which the given system finds itself. The chemical potentials μ_i of the individual components can always be written as functions of the temperature T and pressure p of the system (into which other prescribed system's parameters can always be translated, at least in principle) and the concentrations $x_i = n_i/\sum_j n_j$ of its constituents, that is in terms of intensive characteristics only. It is, however, important to keep in mind that in contrast to the reasoning which led to the Gibbs phase rule pertaining to the case in which reactions in the multiphase system can occur, in typical applications of the equilibrium conditions to concrete chemical reactions one starts with *definite* numbers $n_i^{\rm in}$ of moles of the different constituents (chemical compounds) and then let the reaction(s) occur; the role of the equilibrium conditions - the Callenian concept of virtual states of the system blocked by appropriate constraints (reactions inhibitors) which can froze each of the allowed reactions at an arbitrary stage of its developement and allow thereby to ascribe the thermodynamical potentials to the states of the system with different numbers (allowed by the constraints imposed by the character of the reactions) of moles of the constituents is an important element of the reasoning - in this case being to fix the relations between the final (equilibrium) numbers $n_i = n_i^{\rm eq}$ of moles of the constituents. By the reactions which can take place in the system these final mole numbers are constrained (related to the initial ones) by the relations

$$n_i = n_i^{\rm in} + \sum_{a=1}^{R} \nu_i^{(a)} \zeta^{(a)} \,, \tag{185}$$

in which $\zeta^{(a)}$, is called the **advancement** of the a-th reaction. Together, the equilibrium conditions (184) and the relations (185), perhaps with the conditions of constancy of the total system's energy and of its volume, if the equilibrium is to be reached in an isolated system, should determine the final mole numbers of the chemical compounds involved that is determine all the factors $\zeta^{(a)}$.

With the equilibrium conditions which determine the final state of the system in which chemical reactions occur stated in terms of the chemical potentials one can return to the (already discussed) problem of the (joint) heat effect of a chemical reaction (several reactions occuring simultaneously), that is to the problem of how much heat is absorbed or produced in a transition of a given chemically active system to its final equilibrium state. It has been already established (Lecture II) that if a reaction (or reactions) occur at the fixed temperature T and pressure p, the heat effect is given by the difference of the enthalpies

$$\bar{Q} = -\Delta H \equiv H^{\rm in} - H^{\rm fin} \,.$$

Since $G = H - TS$, the requisite difference of the enthalpies can be easily related to the difference of the Gibbs function of the system:

$$H = G - T\left(\frac{\partial G}{\partial T}\right)_p \equiv -T^2\left(\frac{\partial}{\partial T}\frac{G}{T}\right)_p .$$

Thus

$$\bar{Q} = T^2\left(\frac{\partial}{\partial T}\frac{\Delta G}{T}\right)_p .$$

In turn, since (Lecture VIII)

$$G = \sum_{i=1}^{r} n_i\,\mu_i\,,$$

the infinitesimal change dG of the Gibbs function G due to the infinitesimal advancements $d\zeta^{(a)}$ of the simultaneous R reactions which occur in the system is

$$dG = \sum_{j=1}^{r} dn_j\,\frac{\partial}{\partial n_j}\sum_{i=1}^{r} n_i\,\mu_i = \sum_{j=1}^{r}\mu_j\,dn_j = \sum_{j=1}^{r}\mu_j\sum_{a=1}^{R}\nu_j^{(a)}\,d\zeta^{(a)}\,.$$

(Exploited here has been the fact that $\sum_i n_i(\partial\mu_i/\partial n_j) = 0$ forming the basis of the possibility of treating the system in which chemical reactions can hold as a "black box" and the relation (179)). Since the chemical potentials μ_i of the constituents are in general functions of the concentrations x_i, that is of the actual values of the mole numbers n_i (which in turn are functions of the actual values of the parameters $\zeta^{(a)}$ - see (185)), to obtain the finite change ΔG, the above exact (by the very nature of the Gibbs function G) one-form has to be integrated from the initial values $\zeta^{(a)} = 0$ of the advancements to their final, determined by the equilibrium conditions (184) and the relations (185), values $\zeta_{\rm eq}^{(a)}$ (the exactness of the one-form ensures that the integration path is irrelevant). This task simplifies, however, if one considers reactions in perfect gases.

Chemical reactions in mixtures of (perfect) gases. The mass action law

The equilibrium conditions written above in most cases remain only formal, since the explicit forms of the chemical potentials of the constituents, which are necessary to derive concrete quantitative predictions, are unknown. The simplest case in which at least the pressure and concentration dependence of the chemical potentials can be given explicitly are reactions occuring in perfect gases. In this case (as follows from the Gibbs Ansatz)

$$u_i(T, v_i) = u_i^{(0)} + \int_{T_0}^{T} dT'\,c_v^{(i)}(T')\,,$$

$$s_i(T, v_i) = s_i^{(0)} + \int_{T_0}^{T} dT'\,\frac{c_v^{(i)}(T')}{T'} + R\ln\frac{v_i}{v_i^{(0)}}\,,$$

Therefore the chemical potentials can be written as

$$\mu_i(T, p, x_1, \dots, x_r) = \chi_i(T) + RT\ln p_i \equiv \chi_i(T) + RT\ln p + RT\ln x_i\,, \tag{186}$$

where p_i are the partial pressures of the individual components (perfect gases) constituting the mixture. The exact form of the functions

$$\begin{aligned}\chi_i(T) &= u_0^{(i)} - Ts_0^{(i)} + RT_0 + \int_{T_0}^{T} dT'\,c_p^{(i)} - T\int_{T_0}^{T} dT'\,\frac{c_p^{(i)}}{T'} + RT\ln p_0\\ &\equiv u_0^{(i)} - Ts_0^{(i)} + \gamma^{(i)}(T)\,,\end{aligned}$$

can be computed using statistical physics methods even taking into account possible internal excitations of the individual molecules of the chemical compounds involved, so long the molecules can be treated as mutually noninteracting (Lecture XII). Within pure thermodynamics one is forced to determine these functions with the help of calorimetric measurements. The same is even more true in the case of reactions between substances which cannot be treated as perfect gases - the relevant chemical potentials must be then necessarily determined experimentally; since any calorimetric measurement done between finite temperatures always leaves undetermined constants in $\chi_i(T)$, the complete determination of these functions (necessary to made definite predictions concerning the outcome of reactions) must use an extrapolation to zero temperature where the 3TMDL can be invoked to fix the constants.

We consider first only a single reaction occuring in a mixture of perfect gases. In this case the equlibrium condition

$$\sum_i \nu_i\,\chi_i(T) + RT\sum_i \nu_i\ln p_i = 0\,,$$

is usually rewritten in one of the possible alternative forms either as

$$\prod_{i=1}^{r} p_i^{\nu_i} = \exp\left(-\frac{1}{RT}\sum_{i=1}^{r}\nu_i\,\chi_i(T)\right) \equiv K_p(T)\,, \tag{187}$$

with the right hand side a function of the temperature only, or as

$$\prod_{i=1}^{r} x_i^{\nu_i} = p^{-\sum_i \nu_i}\,K_p(T) \equiv K_x(T, p)\,, \tag{188}$$

or, using the fact that $p_i = (n_i/V)RT$, as the relation between the molar concentrations n_i/V

$$\prod_{i=1}^{r}\left(\frac{n_i}{V}\right)^{\nu_i} = (RT)^{-\sum_i \nu_i}K_p(T) \equiv K_n(T)\,. \tag{189}$$

The first form is called the **mass action law**. It has been established by Guldberg and Waage in 1867. Because in their formulation the temperature dependence of the factor

K_p has been missed, $K_p(T)$ is ever since called the **chemical equilibrium constant** (of the given reaction). Either form of the equilibrium condition can be used to determine the advancement $\zeta_{\rm eq}$ of the reaction obtained after the chemically active system of gases has reached its equilibrium state for given initial conditions (the initial mole numbers of the reacting constituents).

In this particular case the problem of the determination of the heat effect of the reaction, which in general requires integration, simplifies. Indeed, the change dG of the Gibbs function due to the infinitesimal advancement $d\zeta$ of the reaction given by

$$dG = \sum_{i=1}^{r} \mu_i\, dn_i = \sum_{i=1}^{r} \mu_i \nu_i\, d\zeta = \left(RT \sum_{i=1}^{r} \nu_i \ln p_i + \sum_{i=1}^{r} \nu_i\, \chi_i(T) \right) d\zeta\,.$$

takes, after writing the second sum as $-RT\ln K_p(T)$, the form

$$dG = RT\left(\sum_{i=1}^{r} \nu_i \ln(p\, x_i) - \ln K_p(T) \right) d\zeta\,.$$

After dividing by T, the term depending on the concentrations $x_i = x_i(\zeta)$ (varying in the course of reaching the final state) drops out when the resulting expression gets differentiated with respect to T and the coefficient of the one-form

$$dH = -T^2\left(\frac{\partial}{\partial T}\frac{dG}{T} \right) = RT^2\left(\frac{d\ln K_p(T)}{dT} \right) d\zeta\,,$$

is independent of ζ. The form can be therefore immediately integrated from $\zeta = 0$ to $\zeta_{\rm eq}$ determined by the eqilibrium conditions and the heat effect of the reaction is then given by the so called van't Hoff[123] formula

$$\bar{Q} = -RT^2\,\frac{d\ln K_p(T)}{dT}\,\zeta_{\rm eq}\,. \qquad (190)$$

Moreover, since $K_x(T,p)$ given by (188) can be written as

$$K_x(T,p) = \exp\left(-\frac{1}{RT}\sum_{i=1}^{r} \nu_i\, \mu_i(T,p) \right),$$

and $(\partial\mu_i(T,p)/\partial p)_T = v_i$, the change $\Delta V = \sum_i v_i \Delta n_i$ of the volume occupied by the mixture of gases in which the reaction took place is given by

$$\Delta V = -\zeta_{\rm eq}\, RT\left(\frac{\partial \ln K_x(T,p)}{\partial p} \right)_T\,. \qquad (191)$$

---

[123] Do not confuse him with G. 't Hooft - the Nobel Prize winner (together with M. Veltman) for the development of computational techniques for non-Abelian gauge theories and the proof of their renormalizability (whathever the latter notion could mean).

These considerations extend straightforwardly to the case of R reactions occuring simultaneously in the mixture of (perfect) gases. One then has R functions $K_p^{(a)}(T)$ defined by the formulae analogous to (187) and R conditions

$$\prod_{i=1}^{r} p_i^{\nu_i^{(a)}} = \exp\left(-\frac{1}{RT}\sum_{i=1}^{r}\nu_i^{(a)}\chi_i(T)\right) \equiv K_p^{(a)}(T)\,, \qquad a=1,\dots,R\,,$$

which determine the R factors $\zeta^{(a)}$ and the final numbers of moles of the constituents. The differential change of the Gibbs function is given by

$$dG = RT\sum_{a=1}^{R}\left(\sum_{i=1}^{r}\nu_i^{(a)}\ln(p\,x_i) - \ln K_p^{(a)}(T)\right)d\zeta^{(a)}\,,$$

and the finite heat effect is given by

$$\bar{Q} = -RT^2\sum_{a=1}^{R}\frac{d\ln K_p^{(a)}(T)}{dT}\,\zeta_{\rm eq}^{(a)}\,.$$

Phase transitions. Their classification

From the theoretical point of view phase transitions belong to most fascinating phenomena which can be studied with the help of methods of thermodynamics and, primarily, of statistical physics. The subject is vast and gave rise to the developement of important ideas of scaling and methods (which found applications also in high energy physics!) known under the name of the renormalization group. Here we give an introduction to the thermodynamic treatement of phase transitions restricting the discussion to only phase transitions occuring in one-component simple substances. Suppose a quantity (n moles) of such a substance is in a state (characterized by its internal temperature T and pressure p) in which only one of its phases is stable. If this state is changed (by compressing/expanding heating/cooling the system) so that its temperature T and pressure p assume the values at which two (or three) phases can exist, the proportion of the substance in the phases will change - the phase transition will occur.

Two phases can coexist only if their chemical potentials are equal. Thus the chemical potentials are continuos along the coexistence lines. P. Ehrenfest classified phase transitions according to the continuity of derivatives of chemical potentials: in the first order transitions continuous are chemical potentials but their first derivatives (at least one of them) - the molar entropies s and/or molar volumes v, in the case of one component substances - are discontinuous. In the second order phase transitions the chemical potentials and their first derivatives are continuos but at least one of the chemical potentials second derivatives - i.e. one of the derivatives of s or v, that is, c_p, α_p or k_T in the case of one component substances - is discontinuous, and so on.

Of the transitions which fit into this classification only the first order ones are common. Second order transitions are very rare - the main example being the transition conductor-superconductor in zero external magnetic field $\mathcal{H}$ discussed at the end of this Lecture.

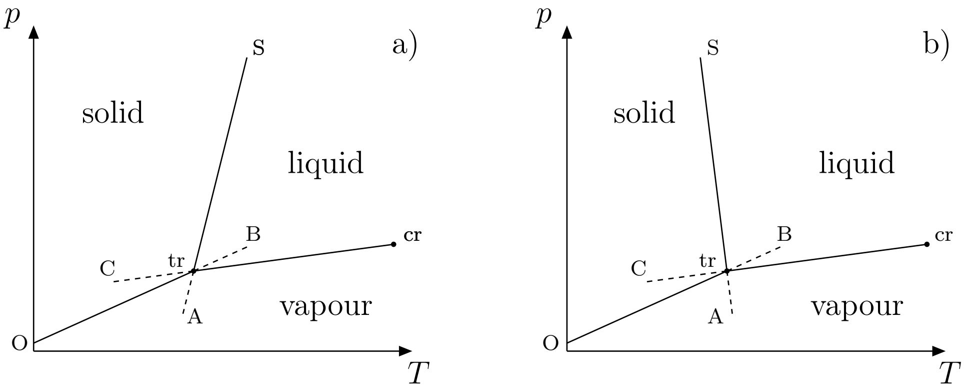

Figure 20: a) Schematic view of a typical phase diagram of a simple one-component substance which in the shown ranges of the temperature and pressure can exist in three phases. b) Schematic view of a slightly less typical phase diagram (H_2O, Bi).

There are, however, many transitions which do not fit the Ehrenfest classification because derivatives of the chemical potentials are divergent at the transition: for instance many first-order phase transitions end up at the critical point at which derivatives become divergent (so it is hard to decide whether they are continuous or not). For this reason Landau simplified the classification: one now distinguishes only the first-order transitions, while all others are called continuous (on account of the continuity of the first derivatives of the chemical potentials).

First-order phase transitions.
We shall now discuss in general terms a typical[124] phase diagram of a simple, single component substance which when conditions are varied, undergoes first-order phase transitions (perhaps excluding from the discussion a few isolated points). The phase diagram of such a substance shows on the (T, p) plane regions in which particular phases are stable. Phase diagrams of multicomponent substances can be considerably more complicated: apart from T and p also molar fractions $x_i^{(\alpha)}$ enter the game and one can only display projections of the stability regions onto various two-dimensional planes.

We consider a simple one-component substance which in the considered ranges of the variables T and p can exist in three modifications (phases): solid, liquid and vapour. Each of the modifications is characterized by its molar Gibbs function $g(T, p)$, that is, its chemical potential $\mu(T, p)$ which can be imagined to exist over the whole (T, p) plane. Thus there are three g-surfaces corresponding to the three phases which pairwise intersect along some lines (along which the two phases can coexist in equilibrium) and there is usually (but not always) one point, called the triple point, at which these three intersections meet - at this point (cf. the Gibbs phase rule) all the three phases can coexist. The projections of these curves onto the (T, p) plane give the phase diagram of the type shown in Figures 20 a and b.

As follows from the discussion carried out at the end of Lecture VII, in a given region of the (T, p) plane stable is that phase which has the lowest value of the molar Gibbs function

---

[124] Atypical phase diagrams will be discusses later.

(because all the substance put in that phase gives the lowest value of the availability A). To determine the regions of stability of different phases one can consider the variations of these functions with T and p which readily follow from the formulae

$$dg \equiv d\mu = -s\,dT + v\,dp\,, \qquad \left(\frac{\partial g}{\partial T}\right)_p = -s\,, \qquad \left(\frac{\partial g}{\partial p}\right)_T = v\,.$$

Thus in the direction of increasing pressure the g-surfaces (μ-surfaces) always slope upwards ($v > 0$): the g-surfaces corresponding to the solid and liquid phases comparatively slowly (small v), at nearly constant rate (small isothermal compressibility) and that of the vapour phase fast (much larger v) with the slope becoming steeper (considerably larger compressibility) in the direction of decreasing pressure. If the arbitrary constant factor in the molar entropy of the substance is fixed so that $s(T,p) \geq 0$, the g-surfaces slope downwards (negative slope) at increasing rate ($c_p > 0$) in the direction of increasing temperature.

From the analysis of the slopes of the g-surfaces in the pressure direction it follows that in the region of low pressures (just above the T-axis) stable is the vapour phase (its Gibbs function falls most steeply as the pressure goes down), but as far as the region near the p axis is concerned, one has to rely on experimental observations to identify the stable phase correctly (^{3}He is an example that the liquid phase can be - in some range of the pressure - closer to the p axis than the solid phase). In the phase diagram shown in Figure 20a crossing the line tr-S vertically upwards one moves from the region in which liquid is the stable phase to the region in which stable is the solid phase; this means that $v^{(\mathrm{solid})} < v^{(\mathrm{liquid})}$: the substance expands on melting. Figure 20b shows the opposite situation in which $v^{(\mathrm{solid})} > v^{(\mathrm{liquid})}$: here the substance (water is here the most obvious example) expands on solidifying. When a coexistence line in Figure 20a or 20b is crossed horizontally (at constant pressure) the phase which is stable on the right hand side must have larger entropy, which means that in Figure 20 in the transitions solid→liquid, solid→vapour or liquid→vapour heat is absorbed. Finite difference of entropies of adjacent phases means that if the system is in the lower entropy phase and heat is supplied to it at a constant rate (at constant pressure, say), its temperature rises until the coexistence line is reached; then there is a halting of the temperature rise: the new phase appears and the amount of substance in it grows. Only when the phase transition is accomplished - there is only the new phase - does the temperature begin to rise again (the same is observed when the heat is supplied and the volume of the system is simultaneously changed in such a way that the coexistence line is crossed not exactly in the horizontal direction). The halting of the temperature grow is the distinguishing feature of first order phase transitions. The heat absorbed by the system in the transition is called the **latent heat**.

Even if in each of the regions (with the exception of the coexistence curves) only one phase is stable, it makes sense to view the g-surfaces as extending beyond the regions of stability of the corresponding phases because from the experience it is known that the phases (particularly if the substances are highly purified) can exist as metastable states in the regions in which they should have already turned into another phase. Supercooled vapours can stay uncondensed at pressures significantly higher than the equilibrium vapour pressure (by which term one understands the pressure at the vapour-liquid coexistence curve at a given temperature) - the historic device of particle physics, the Wilson cloud chamber, owed its operation to this phenomenon), purified liquids can be supercooled without solidifying - droplets of water in a cloud may be cooled down to -40^oC. And conversely, liquids can be superheated - another historic device of particle physics, the bubble chamber, exploited this possibility. Finally it should be mentioned that also the coexistence lines can be prolonged into the regions in which none of the two coexisting along it phases is stable as indicated by the dashed lines in Figure 20 - coexistence of supercooled liquid with its vapour is well known to meteorology (though it seems a solid-liquid coexistence in the vapour region has never been observed). It should be however kept in mind that the analogous continuation of the g-surfaces is not necessarily justified in the case of higher order transitions.

The Clapeyron-Clausius equation

The coexistence curve of two phases of a simple single-component substance is determined by the equality of the two molar Gibbs functions (chemical potentials) and if these were known precisely, the form of the coexistence curve would be also known. But since the chemical potentials as functions of T and p are usually not given, the Clapeyron-Clausius equation allows to reconstruct the coexistence curve on the basis of the more directly accessible experimental data. It is derived straightforwardly by taking the total derivative with respect to T of the equality

$$\mu^{(\alpha)}(T,\, p(T)) = \mu^{(\beta)}(T,\, p(T))\,.$$

Another way of deriving this equation (better adapted to higher order transitions) consists of expanding in the Taylor series both sides of the equality

$$\mu^{(\alpha)}(T+\Delta T,\, p+\Delta p) = \mu^{(\beta)}(T+\Delta T,\, p+\Delta p)\,.$$

Either way, if at least one of the two derivatives of the chemical potential is discontinuous one obtains

$$\frac{dp}{dT} = \frac{s^{(\beta)} - s^{(\alpha)}}{v^{(\beta)} - v^{(\alpha)}} = \frac{q_{\alpha\to\beta}}{T(v^{(\beta)} - v^{(\alpha)})}\,, \tag{192}$$

where $T(s^{(\beta)} - s^{(\alpha)})$ has been identified with the latent heat $q_{\alpha\to\beta}$ of the transition. If all quantities on the right-hand side of this equation are given as functions of T and p, the equation can be integrated. One usually makes some crude approximations: eg. in considering the solid-vapour transition the solid molar volume is neglected, the vapour is treated as a perfect gas, $q_{\alpha\to\beta}$ is assumed to be constant etc. (This will be done in classes.)

The Clapeyron-Clausius (or, more familiarly, Clapaucius) equation clearly shows the discussed already correlation of the sign of the slope of the melting curve with the difference of the molar volumes (or densities $\rho = m_{\rm mol}/v$) of the solid and liquid phases. Water molar volume is smaller than that of ice (which has also lower entropy - the latent heat

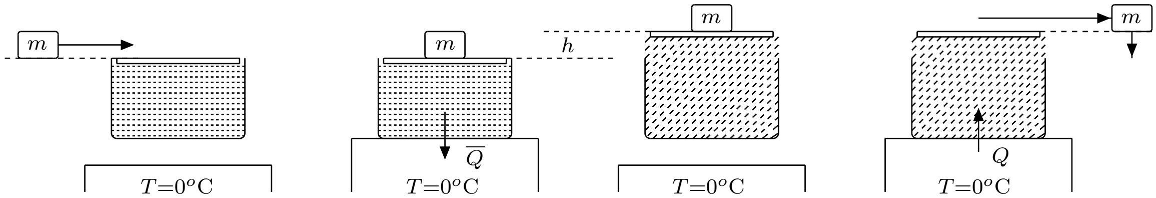

Figure 21: The (impossible) "ice-water" engine violating 2TMDL.

$q_{\alpha\to\beta}$ is defined to be positive, so $\alpha =$ice, $\beta =$water) which results in the negative slope (negative right hand side of the equation (192)) of the ice melting curve (which is of the type shown in Figure 20b). That the slope of the ice melting curve cannot be infinite (the curve cannot be vertical, that is, the melting temperature of ice cannot be independent of pressure, given that ice has larger molar volume than water) at least in the range of pressures in which the ice molar volume is greater than that of the water, was understood by the second Thomson, the lord Kelvin's brother, around the middle of the XIX century to be required by 2TMDL. If the melting curve were vertical, one could devise an "ice-engine", shown in figure 21, which would violate the Kelvin's 2TMDL: starting with a quantity of water at 0^{o}C one could place a weight m on the piston enclosing the vessel with the water from above (the water compressibility is small and can be neglected here). Then a quantity $\bar{Q}$ of heat could be extracted from the water to a reservoir at 0^{o}C, as a result of which the water would solidify increasing its volume (maintaining its initial temperature, the heat extracted from it being the latent heat) and raising thereby the weight to some height $h > 0$; the weight could be then moved to the side and brought down to its initial height delivering some mechanical work (at the cost of its gained potential energy), while an appropriate quantity Q of heat could be supplied (again from the reservoir at 0^{o}C) to the ice causing it all to melt; in this way the working substance - the water - would be in its initial state and the work $\bar{W} = mgh$ would be done at the cost of the heat $Q - \bar{Q}$ taken entirely from the single reservoir at 0^{o}C. This is something the Kelvin's 2TMDL forbids. The resolution is of course that in the proposed cycle the processes of solidifying water and of ice melting occur at different pressures - solidifying occurs at $p = p_0 + mg/A$, while melting at $p = p_0$, where p_0 is the pressure of the surrounding and A is the section area of the piston - and 2TMDL is saved if the melting/solidifying temperature varies appropriately with the pressure. The comparison of the measured pressure variation of the ice melting point with that predicted by the Clapaucius equation was one of the first successful applications of thermodynamics to physical problems and significantly enhanced the confidence in this developing branch of theoretical physics.

Critical point

The line of the coexistence of the liquid and vapour phases (the line tr-cr on the diagrams 20a,b) does not continue indefinitely: as the line is followed to the right, the quantitative characteristics (densities, heat capacities, compressibilities, etc.) of the two phases: the vapour and the liquid ones, become more and more similar and all differences between them eventually disappear at the **critical point** marked "cr" (and characterized by $T_{\rm cr}$, $p_{\rm cr}$ and, through the equation of state, also by $v_{\rm cr}$) at which the liquid and vapour become indistinguishable.[125] To discuss the nature of the critical point it is convenient to plot several isotherms of the vapour-liquid system on the (v, p) indicator diagram.

If, by decreasing its volume, the vapour at $T < T_{\rm cr}$ is compressed isothermally (one moves vertically upwards on the phase diagrams 20a,b) the isothermal compressibility $k_T = -(\partial \ln v/\partial p)_T$ first (far below the coexistence line) decreases and then rises to become infinite at the coexistence line, when the liquid begins to form; the system becomes there inhomogeneous breaking up into two separate phases; subsequent decreasing its volume does not increase the pressure (k_T stays infinite), it only increases the quantity of the liquid as compared to the quantity of the vapour (on the diagram 20 one is staying all this time on the coexistence curve); when the vapour has disappeared completely, the pressure begins to rise rather steeply (k_T falls abruptly to a small value) and one moves on the diagram 20 upwards already above the coexistence line.

When the analogous compression of the vapour is effected exactly at the critical temperature $T_{\rm cr}$, the compressibility k_T first decreases as the volume is reduced (and pressure rises) then rises to become momentarily infinite at $p_{\rm cr}$ and then again steadily goes down. At no point in this case is a separation of the system into two phases observed. Above $T_{\rm cr}$ as the volume is reduced, the compressibility decreases monotonically: at low values of the pressure (large volume) it is large so the substance resembles more the vapour, while at high pressures (smaller volumes) it is small, so the system behaves more as a liquid.

It follows that it is possible by traveling around the point "cr" on the diagram 20 to make the system to pass from the vapour phase to the liquid phase, without any discontinuous changes in its properties.[126] This means that the two g-surfaces: the $g_{\rm l}(T, p)$ one and the $g_{\rm v}(T, p)$ one which we compared in establishing which of the phases is stable in the given domain of the (T, p) plane, are in fact two parts of one and the same g-surface which for $T < T_{\rm cr}$ simply intersects itself along a line the projection of which onto the (T, p) plane is the curve tr-cr, but not for $T > T_{\rm cr}$.

One way of visualizing the form of such a self-intersecting g-surfaces is to use the Van der Waals equation of state

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT\,, \tag{193}$$

---

[125] The existence of critical points has been first noticed some 150 years ago in connection with the problem of liquefaction of gases: some gases seemed impossible to liquify by applying even the highest available pressures. It is now clear that this was because one attempted to liquify them at temperatures higher than their critical temperatures. These can be very low indeed: the critical temperature of He is only 5.2 K. It should be also recalled that polish physicists Z.F. Wróblewski and K.S. Olszewski succeeded in 1883 to liquify for the first time O_2, then N_2, Air and CO by cooling them down to 113 K; later Olszewski liqified also Ar by achieving $T = 48$ K but this was still not low enough to allow for liquefaction of H_2.

[126] It is appropriate to make a comment on the terminology: in the older literature the system at temperatures $T < T_{\rm cr}$ is called either liquid (at higher pressures) or vapour (at lower pressures) while the system above $T_{\rm cr}$ is called gas. More adequate seems to call the system above $T_{\rm cr}$ supercritical fluid and the one below $T_{\rm cr}$ liquid and vapour or, exchangeably, gas.

which qualitatively (though not very well quantitatively) models the equation of state of real vapour-liquid systems. High T isotherms given by this equation are perfectly monotonic and the relation of the pressure p to the molar volume is one-to-one, just as happens with the $T > T_{\rm cr}$ isotherms of a real vapour-liquid system. Below some temperature, which should therefore be identified with $T_{\rm cr}$, to a given pressure p there correspond three different values (of which the smallest one and the largest one we, introducing thereby their interpretation, denote $v_{\rm l}$ and $v_{\rm v}$, respectively) of the molar volume. It can be also seen that the inverse compressibility k_T^{-1}, which can be calculated, given the equation of state,

$$k_T^{-1} = -v\left(\frac{\partial p}{\partial v}\right)_T = \frac{RTv}{(v-b)^2} - \frac{2a}{v^2}\,,$$

is always positive, if $T > T_{\rm cr}$, but for $T < T_{\rm cr}$ it becomes negative somewhere between $v_{\rm l}$ and $v_{\rm v}$ (therefore the compressibility k_T is there negative too). This means that the intrinsic stability condition (discussed in Lecture VII) is violated there and one is forced to admit that these parts of subcritical VdW isotherms do not represent system's states (are unphysical). In agreement with what happens in real systems, some parts of the subcritical VdW isotherms should be therefore replaced by flat horizontal lines, because one knows that as the coexistence curve is reached on the diagram 20 from below (or from above), the system breaks up into two phases and its number of moles n and the molar volume v are given by

$$n = x_{\rm l}n_{\rm l} + x_{\rm v}n_{\rm v}\,, \qquad v = x_{\rm l}v_{\rm l} + x_{\rm v}v_{\rm v} \qquad x_{\rm l} + x_{\rm v} = 1\,,$$

with $x_{\rm v}$ beginning at 1 (at 0) and decreasing to 0 (increasing to 1) as the total volume is reduced (increased).

The possible interpretation of the subcritical VdW isotherms is therefore as follows (Figure to be done - the position of the horizontal line will be discussed below):

- the segment at large v going through the point A up to the point B correspond to the vapour (gas);
- point B is where the phase separation should begin (it corresponds to the reaching the line tr-cr on the diagram 20 from below);
- the segment B-C can be interpreted as representing the supercooled vapour (a metastable state) - on Figure 20 the corresponding points lie above the tr-cr line;
- the segment C-D-E is unphysical ($k_T < 0$);
- the segment E-F can be interpreted as representing the superheated liquid (a metastable state) - on Figure 20 the corresponding points lie below the tr-cr line;
- the flat segment B-D-F represents the mixture of phases in neutral equilibrium

- the segment from the point F through G and further to the left represents the system in the single liquid phase

Having the VdW equation one can consider integrating up the one-form

$$dg = -s(T,p)\,dT + v(T,p)\,dp\,,$$

along the shown isotherm $T < T_{\rm cr}$ from the point, say A, on:[127]

$$g(T,p) = g(T,p_A) + \int_{p_A}^{p} dp'\,v(T,p')\,. \tag{194}$$

As we integrate, the area under the curve $v = v(p)$ (at fixed T) grows until the point C is reached, so the plot of the function $g(T,p)$ raises. Then from C to E, as p decreases, the area must be counted as negative; therefore the plot of $g(T,p)$ goes backwards and falls down. Finally, from the point E on p again increases and the area under the curve $v = v(p)$ must be counted as positive, so the plot of $g(T,p)$ again goes forward and raises. The resulting self crossing curve $g(T,p)$ (a $T =$const section of the self-crossing g-surface) is shown in Figure. Its part A-C is the constant T section of what formerly was supposed to be the vapour g-surface, while its part E-G is the section of the liquid g-surface; up to the point at which they cross the vapour g-surface is lower and it is vapour which is the stable phase, while from this point on stable is the liquid phase. As the temperature increases, the unphysical segment C-E becomes shorter and shorter and disappears completely at $T_{\rm cr}$; making plots of several such self crossing g-curves one obtains the picture of the self crossing g-surface of the liquid vapour system.

There remains only the question where to place the points B and F on a given isotherm? The prescription given by Maxwell (and called the Maxwell construction ever since) states that these points should be drawn in such a way that the areas B-C-D-B and D-E-F-D be equal. In this way the value of the molar Gibbs function at the point B will be the same as the value of the molar Gibbs function at the point F (that is, equal will be the values of the chemical potentials of the vapour and of its liquid at the temperature T labeling this isotherm). Indeed, the Maxwell rule requires that (geometrically)

$$(\text{area under B} - \text{C}) - (\text{area under C} - \text{D}) = (\text{area under D} - \text{E}) - (\text{area under E} - \text{F})$$

But this precisely ensures that (recall how the integration (194) has been done)

$$\begin{aligned} g(T,p_F) - g(T,p_B) &= \int_{p_B}^{p_F} dp\,v(T,p) \\ &= (\text{area under B} - \text{C}) - (\text{area under C} - \text{D}) \\ &\quad - (\text{area under D} - \text{E}) + (\text{area under E} - \text{F}) = 0\,. \end{aligned} \tag{195}$$

---

[127] Of course the parametrization of the isotherm by the pressure is only piecewise well defined and has to be changed on the way - in the discussion below this is taken into account in the "school way", by counting the areas under the curve as positive or negative.

The construction is however of doubtful validity for it clearly exploits the unphysical part of the isotherm (its segment C-E). The correct procedure would be to start from the point A and move on along the isotherm computing at every point the chemical potential just using the standard thermodynamic data including the equation of state - the VdW one in the discussed case - and at the same time to start form the point G and move back along the same isotherm, again computing at every point the chemical potential. In this way one could in principle find the two points, one on the left part of the isotherm and the other one on its right part, at which the two chemical potentials (molar Gibbs functions) have the same value at the same pressure. (This procedure does not exploit the unphysical part of the isotherm). It would be these points, therefore, which should be connected by the flat segment. It may be argued that this shows that the Maxwell construction cannot be correct because if the isotherm between the points B and F is distorted, an operation which would not alter the equality of the chemical potentials at B and F found in the way proposed here, the two areas compared in the Maxwell construction would be modified and it would predict incorrectly the position of the flat part of the isotherm. This argument can be, however, dismissed by appealing to the analyticity property of thermodynamical potentials (everywhere with the exception of the critical point which is a bifurcation point): the "stiffness" of analytic functions does not allow for distorting the unphysical parts of the isotherm in a way uncorrelated with the changes of the chemical potentials found by the "move and control" method.

Critical exponents

Besides the isothermal compressibility k_T in the vicinity of the critical point also other quantities exhibit anomalies. This can be illustrated by modeling again the liquid-vapour system by the VdW equation of state. The molar specific heat c_v (at constant molar volume v) is predicted by this equation to be independent of v and, therefore one can tentatively[128] take it to be constant in the vicinity of the critical point. The relation

$$c_p = c_v + T\left(\frac{\partial p}{\partial T}\right)_v\left(\frac{\partial v}{\partial T}\right)_p = c_v + Tv\,k_T\left(\frac{\partial p}{\partial T}\right)_v^2, \tag{196}$$

then implies that the molar heat capacity at constant pressure should diverge at the critical point because the isothermal compressibility k_T does, as has been already discussed. To capture the character of this divergence it is convenient to rewrite the VdW equation in terms of the **reduced variables** $\pi \equiv p/p_{\rm cr}$, $\omega \equiv v/v_{\rm cr}$ and $t \equiv T/T_{\rm cr}$ (the values $v_{\rm cr} = 3b$, $p_{\rm cr} = a/27b^2$, $RT_{\rm cr} = 8a/27b$ are the solutions of the conditions $(\partial p/\partial v)_T = 0$, $(\partial^2 p/\partial v^2)_T = 0$).

$$\left(\pi + \frac{3}{\omega^2}\right)(3\omega - 1) = 8t\,,$$

and then to go over to the variables $x = \omega - 1$, $y = \pi - 1$ and $\tau = t - 1$ which directly measure the departures from the critical point. Expanded up to the terms of order x^3 the

---

[128] c_v of real fluids can in fact exhibit anomalous behaviour.

VdW equation takes then the simple form

$$3x^3 + (2 + 3x)y - 8\tau = 0\,. \tag{197}$$

The derivatives appearing in the relation (196) can be then straightforwardly computed to yield

$$\left(\frac{\partial p}{\partial T}\right)_v \propto \left(\frac{\partial y}{\partial \tau}\right)_x = \frac{8}{2+3x}\,, \qquad \left(\frac{\partial v}{\partial T}\right)_p \propto \left(\frac{\partial x}{\partial \tau}\right)_y = \frac{8}{3(3x^2+y)}\,,$$

and one finds that (including all numerical factors)

$$c_p = c_v + \frac{8Rt}{(2+3x)(3x^2+y)}\,.$$

It then follows that if $p = p_{\rm cr}$, i.e. when $y = 0$, $c_p \sim {\rm const.}/x^2$, as $x \to 0$. Since from (197) $\tau \propto x^3$ when $y = 0$, it follows that

$$c_p \sim \frac{\rm const.}{|T - T_{\rm cr}|^{2/3}}\,, \tag{198}$$

as $T \to T_{\rm cr}$ at $p = p_{\rm cr}$, whereas if $x = 0$, $c_p \sim {\rm const.}/y$, that is (because $\tau \propto y$ at $x = 0$)

$$c_p \sim \frac{\rm const.}{T - T_{\rm cr}}\,, \tag{199}$$

as $T \to T^+_{\rm cr}$ at $v = v_{\rm cr}$ (this can be computed only for $T > T_{\rm cr}$, for at $v = v_{\rm cr}$ and $T < T_{\rm cr}$ one is precisely in the region of parameters in which the VdW equation leads to unphysical results and one should rather ask about the behaviour of $c_p^{\rm liq}$ and $c_p^{\rm vap}$ as one approches the critical point along the coexistence line). Using the relation $k_S = (c_v/c_p)k_T$ it can be also shown that the adiathermal compressibility k_S remains finite: since (again using the VdW equation in the form (197)) near the critical point

$$k_T \equiv -\frac{1}{v}\left(\frac{\partial v}{\partial p}\right)_T = -\frac{1}{p_{\rm cr}(1+x)}\left(\frac{\partial x}{\partial y}\right)_\tau = \frac{1}{p_{\rm cr}(1+x)}\,\frac{2+3x}{3(3x^2+y)}\,,$$

one gets that $k_T/c_p \propto (2 + 3x)^2(1 + x)^{-1}$.

In general, singularities[129] at the critical point of various quantities characterizing different physical system turn out to be power-like and are conveniently quantified by the so called **critical exponents** which can be defined as follows. Suppose one approaches the critical point along a well defined path which does not coincide with the critical isotherm. Parametrizing this path with $\tau \equiv (T - T_{\rm cr})/T_{\rm cr}$ one defines the exponent λ characterizig a quantity X by

$$\lambda = -\lim_{\tau\to 0}\frac{d\ln X(\tau)}{d\ln\tau}\,. \tag{200}$$

---

[129] By “singularity” of a quantitity one understands either its vanishing or diverging.

One then has[130]

$$X(\tau) = \begin{cases} A\,\tau^{-\lambda} & T > T_{\rm cr}, \\ A'\,(-\tau)^{-\lambda'} & T < T_{\rm cr}, \end{cases} \tag{201}$$

In the similar way one can introduce exponents characterizing approaching the critical point along paths parametrized by variables other than τ. Of course, critical exponents characterize only the leading behaviour of $X(\tau)$ near the critical point: in most cases it is more fully captured by the formula $X(\tau) = A\,\tau^{-\lambda}(1 + B\,\tau^{\kappa} + \ldots)$ with $\kappa > 0$. Interestingly, in most cases the critical exponents on both sides (if both can be defined) are equal: $\lambda = \lambda'$, but $A \neq A'$. A fluid near its critical point:

$$\begin{aligned} & v_{\rm l} - v_{\rm v} = {\rm const.} \times (-\tau)^{\beta}\,, \\ & k_T = \begin{cases} A\,\tau^{\gamma} & T > T_{\rm cr}, \quad v = v_{\rm cr} \\ A'\,(-\tau)^{\gamma'} & T < T_{\rm cr}, \quad v = v_{\rm v} \ \text{ or } \ v = v_{\rm l} \end{cases} \\ & p - p_{\rm cr} = {\rm const.} \times |v - v_{\rm cr}|^{\delta}{\rm sgn}(v - v_{\rm cr})\,. \end{aligned} \tag{202}$$

The VdW equation gives (classes) $\delta = 3$ and $\beta = 1/2$ but this is not what is measured. The precise computation of the critical exponents requires introducing (within the statistical physics) essentially new ideas and methods known under the name of the renormalization group methods.

The conductor-superconductor phase transition

We will briefly consider here the conductor-superconductor phase transition with the aim of showing that in the zero external magnetic field it is of the second order according to the Ehrenfest classification.

Many metals (which do not exhibit ferromagnetic properties) like tin (Sn), iridium (Ir), niobium (Nb), become superconducting at low temperatures (in zero magnetic field at 3.73 K, 0.14 K and 9 K, in the case of tin, iridium and niobium, respectively - we will not consider high temperature superconductors here), that is exhibit zero electric resistance. This is accompanied, and one can show that vanishing of the resistance is a consequence of this effect, by their perfect diamagnetism: if they are placed in an external magnetic field, the magnetic field induction is expelled from their interior altogether (i.e. $\boldsymbol{\mathcal{B}} = \mathbf{0}$ inside the specimen; this is called the Meissner-Ochsenfeld effect) unless the external field is too strong: if it exceeds some critical value $\mathcal{H}_{\rm c}$ (which depends on the temperature, and to a much smaller extent, on the pressure), the material becomes "normal": the resistivity reappears and the magnetic field induction $\boldsymbol{\mathcal{B}}$ in the material is related to the strength $\boldsymbol{\mathcal{H}}$ inside the specimen by a magnetic sussceptibility χ: $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{H}} + 4\pi\boldsymbol{\mathcal{M}}$ where $\boldsymbol{\mathcal{M}} = \chi\boldsymbol{\mathcal{H}}$. Below we will assume that the normal phase sussceptibility χ is negligibly small, so that in the normal state simply $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{H}}$ inside the material (just as outside it).

Suppose a tin specimen of a long cylindrical shape is placed in a weak external magnetic field[131] $\mathcal{H}$ at a temperature lower than the critical one, $T_{\rm c}$ (so that the specimen is in the superconducting phase). The magnetic field induction inside the specimen is

---

[130] $\lambda = 0$ corresponds to a logarithmic singularity.

[131] Since we consider the magnetic field aligned always in the same direction, we can play with the scalar

then zero. If $\mathcal{H}$ is slowly increased so that it exceeds some critical value $\mathcal{H}_{\rm c}$, the metal undergoes a first order (as will be seen) phase transition to the normal state in which $\mathcal{B} = \mathcal{H}$ inside it. If the external field is reduced back below the critical value $\mathcal{H}_{\rm c}$, the material reverts to the perfectly diamagnetic state. The transition is therefore reversible in the thermodynamic sense and can be analyzed with standard methods. The experimentally measured dependence of the critical field strength $\mathcal{H}_{\rm c}$ on the temperature (at fixed pressure) is conveniently modeled by the empirical formula

$$\mathcal{H}_{\rm c}(T) = \mathcal{H}_0 \left[1 - \left(\frac{T}{T_{\rm c}}\right)^2\right], \tag{203}$$

which correctly captures the two essential features of the real dependence: its flatness at $T = 0$ (required by 3TMDL - of which we had no time to talk) and the finite slope at $T = T_{\rm c}$ (which is crucial for further considerations). The dependence on the pressure is rather weak: $T_{\rm c}$ of tin changes from 3.73 K at $p = 1$ atm to 3.63 K at $p = 1700$ atm.

It is convenient to consider the molar magnetic Gibbs function of the metal which exhibits superconductivity

$$g'(T, p, \mathcal{H}) = u' - Ts + p\,v - \mathcal{H}\mathcal{M}v\,, \tag{204}$$

(we assume that the the magnetic field strength $\mathcal{H}$ and the magnetization $\mathcal{M}$ are constant in the material) the differential $dg' = -s\,dT + v\,dp - v\mathcal{M}\,d\mathcal{H}$ of which implies that

$$\left(\frac{\partial g'}{\partial T}\right)_{p,\mathcal{H}} = -s\,, \qquad \left(\frac{\partial g'}{\partial p}\right)_{T,\mathcal{H}} = v\,, \qquad \left(\frac{\partial g'}{\partial \mathcal{H}}\right)_{T,p} = -v\mathcal{M}\,, \tag{205}$$

and which, being a function of intensive parameters only, plays the role of the chemical potential.

Since the system is characterized by three intensive parameters, its two phases, the normal one denoted n and the superconducting one denoted s, in the three dimensional space can, as follows from the Gibbs phase rule, coexist along a two dimensional surface at which their molar gibbs functions are equal $g'_{\rm s} = g'_{\rm n}$. The coexistence surface projected onto the three planes: $(T, \mathcal{H})$, $(p, \mathcal{H})$ and the (T, p) one, gives rise to three coexistence curves which are determined by the equations analogous to the Clapaucius one - the method of derivation is exactly the same:[132]

$$\left(\frac{\partial \mathcal{H}_{\rm c}}{\partial T}\right)_p = -\frac{s_{\rm n} - s_{\rm s}}{v_{\rm n}\mathcal{M}_{\rm n} - v_{\rm s}\mathcal{M}_{\rm s}}\,,$$

---

quantity $\mathcal{H}$. Moreover, because of the assumed specimen shape, the magnetic field strength $\boldsymbol{\mathcal{H}}$ inside the material is the same as the strength $\boldsymbol{\mathcal{H}}_0$ of the applied field (the demagnetization field is zero). Hence we do not need to distinguish them. The symbol $\mathcal{H}_0$ will be therefore used to denote the critical value of the magnetic firld strength at $T = 0$.

[132] E.g. if the pressure is fixed, on the coexistence curve the changes $d\mathcal{H}_{\rm c}$ and dT are correlated so that $g'_{\rm n}(T + dT,\ \mathcal{H}_{\rm c} + d\mathcal{H}_{\rm c},\ p) = g'_{\rm s}(T + dT,\ \mathcal{H}_{\rm c} + d\mathcal{H}_{\rm c},\ p)$.

$$\left(\frac{\partial \mathcal{H}_{\rm c}}{\partial p}\right)_T = \frac{v_{\rm n} - v_{\rm s}}{v_{\rm n}\mathcal{M}_{\rm n} - v_{\rm s}\mathcal{M}_{\rm s}}\,, \tag{206}$$
$$\left(\frac{\partial p}{\partial T}\right)_{\mathcal{H}} = \frac{s_{\rm n} - s_{\rm s}}{v_{\rm n} - v_{\rm s}}\,.$$

As in the superconducting (perfectly diamagnetic) phase $\mathcal{B} = 0$ means that $\mathcal{M}_{\rm s} = -\mathcal{H}/4\pi$, while the equality $\mathcal{B} = \mathcal{H}$ holding in the normal phase implies that in this phase $\mathcal{M}_{\rm n} = 0$, the first two of the above three equations can be rewritten in the forms

$$\left(\frac{\partial \mathcal{H}_{\rm c}}{\partial T}\right)_p = -\frac{4\pi}{v_{\rm s}\mathcal{H}_{\rm c}}\,(s_{\rm n} - s_{\rm s})\,, \qquad \left(\frac{\partial \mathcal{H}_{\rm c}}{\partial p}\right)_T = \frac{4\pi}{v_{\rm s}\mathcal{H}_{\rm c}}\,(v_{\rm n} - v_{\rm s})\,. \tag{207}$$

The values $s_{\rm s}$, $s_{\rm n}$, etc. should be taken on the respective coexistence surfaces, but their variation with the magnetic field is not large: since, as readily follows from (205), $(\partial s/\partial\mathcal{H})_{T,p} = (\partial v\mathcal{M}/\partial T)_{p,\mathcal{H}}$, neither $s_{\rm n}$ nor $s_{\rm s}$ are appreciably field-dependent: the former because $\mathcal{M}_{\rm n} = 0$ and the latter because $\mathcal{M}_{\rm s} = -\mathcal{H}/4\pi$, so only a rather small temperature dependence of $v_{\rm s}$ can contribute to the variation of $s_{\rm s}$ with the magnetic field strenhth. Similarly, since $(\partial v/\partial\mathcal{H})_{T,p} = -(\partial v\mathcal{M}/\partial p)_{T,\mathcal{H}}$, $v_{\rm n}$ is field independent and to the field dependence of $v_{\rm s}$ can contribute only the small compressibility of the specimen in this phase. Thus, with the exception of the differences $s_{\rm n} - s_{\rm s}$ and $v_{\rm n} - v_{\rm s}$ one can use in the above equations $s_{\rm s}$ and $v_{\rm s}$ taken at the zero magnetic field and similarly neglect the field dependence of $s_{\rm n}$ and $v_{\rm n}$.

Now, the first one of the two equations (207) rewritten in the form

$$s_{\rm n} - s_{\rm s} = -\frac{v_{\rm s}\mathcal{H}_{\rm c}}{4\pi}\left(\frac{\partial \mathcal{H}_{\rm c}}{\partial T}\right)_p\,, \qquad v_{\rm n} - v_{\rm s} = \frac{v_{\rm s}\mathcal{H}_{\rm c}}{4\pi}\left(\frac{\partial \mathcal{H}_{\rm c}}{\partial p}\right)_T\,, \tag{208}$$

shows, because - as already has been stressed and as the formula (203) makes it clear - the slope $(\partial\mathcal{H}_{\rm c}/\partial T)_p$ is not infinite at $\mathcal{H}_{\rm c} = 0$, that at $\mathcal{H}_{\rm c} = 0$ the difference $s_{\rm n} - s_{\rm s}$ vanishes. This means that in zero magnetic field the transition conductor-superconductor occurs without any latent heat.[133] In the same way the second equation shows that in zero magnetic field vanishes also the difference $v_{\rm n} - v_{\rm s}$ (because $\mathcal{H}_c$ depends on p very weakly and the derivative is certainly not infinite). Thus, in the ordinary two-parameter space (T, p) corresponding to zero magnetic field, the conductor-superconductor transition ceases to be first order. This means that the third of the equations (206) is indeterminate at zero field.[134] To show that the transition is of second order according to the Ehrenfest classification, it is sufficient to show that the derivatives of entropies and/or of molar volumes are discontinuous across the coexistence curve. Differentiating the first equation

---

[133] The same hold also at $T = 0$, because $(\partial\mathcal{H}_{\rm c}/\partial T)_p = 0$ there; this is consistent with 3TMDL which requires equality of molar entropies of both phases to be equal at absolute zero and therefore forbids existence of first order phase transitions at this temperature.

[134] The transition remains first order in all two-parameter spaces (T, p) corresponding to non-zero magnetic field and on these planes the third equation (207) is well defined and does determine the coexistence curve.

(208) with respect to T at constant pressure (again neglecting the temperature dependence of $v_{\rm s}$ in their right hand sides) one finds

$$c_p^{(\rm n)} - c_p^{(\rm s)} = -\frac{v_{\rm s}T}{4\pi}\left[\mathcal{H}_c\left(\frac{\partial^2\mathcal{H}_{\rm c}}{\partial T^2}\right)_p + \left(\frac{\partial\mathcal{H}_{\rm c}}{\partial T}\right)_p^2\right] = -\frac{v_{\rm s}T}{8\pi}\left(\frac{\partial^2\mathcal{H}_{\rm c}^2}{\partial T^2}\right)_p . \tag{209}$$

Appealing to the empirical formula (203) one sees that at zero magnetic field the discontinuity of the molar heat capacities is finite and is equal[135]

$$c_p^{(\rm n)} - c_p^{(\rm s)} = -\frac{v_{\rm s}\mathcal{H}_0^2}{\pi T_{\rm c}} .$$

Since the discontinuity of molar heat capacities of the two phases is finite, the transition is of second order. It remains to see what replaces the third, indeterminate equation (206) in determining the conductor-superconductor coexistence curve on the (T, p) plane (at zero magnetic field). To this end one takes the derivatives at $\mathcal{H} = 0$ of both equalities (208) with respect to T and with respect to p (again neglecting any small dependence of $v_{\rm s}$ on these parameters) obtaining three relations (the fourth one is identical with one of these three, because of the Maxwell identity $(\partial s/\partial p)_T = -(\partial v/\partial T)_p$) which can be written in the forms:[136]

$$\begin{aligned}
\left(\frac{\partial\mathcal{H}_{\rm c}}{\partial T}\right)_p^2 &= -\frac{4\pi}{vT}\left(c_p^{(\rm n)} - c_p^{(\rm s)}\right) ,\\
\left(\frac{\partial\mathcal{H}_{\rm c}}{\partial p}\right)_T^2 &= -4\pi\left(k_T^{(\rm n)} - k_T^{(\rm s)}\right) ,\\
\left(\frac{\partial\mathcal{H}_{\rm c}}{\partial T}\right)_p\left(\frac{\partial\mathcal{H}_c}{\partial p}\right)_T &= 4\pi\left(\alpha_p^{(\rm n)} - \alpha_p^{(\rm s)}\right) .
\end{aligned}$$

Taking their ratios and exploiting the standard shocking relation $(\partial\mathcal{H}_{\rm c}/\partial p)_T(\partial\mathcal{H}_{\rm c}/\partial T)_p = -(\partial T_{\rm c}/\partial p)_{\mathcal{H}}$, one obtains the two (mutually consistent, as follows from the construction) equations

$$\left(\frac{\partial T_{\rm c}}{\partial p}\right)_{\mathcal{H}=0} = v_{\rm s}T\,\frac{\alpha_p^{(\rm n)} - \alpha_p^{(\rm s)}}{c_p^{(\rm n)} - c_p^{(\rm s)}} = \frac{k_T^{(\rm n)} - k_T^{(\rm s)}}{\alpha_p^{(\rm n)} - \alpha_p^{(\rm s)}} ,$$

which replace the last of the equations (206) in determining the second order phase transition line on the (T, p) plane. The equation obtained here, called the Ehrenfest equation, is the analog of the Clapeyron-Clausius equation in the case of second order phase transitions.

---

[135] Since we restrict our attention to the zero magnetic field plane, we denote these heat capacities c_p; beyond that plane one should call it more properly $c_{p,\mathcal{H}}$.

[136] Recall that $k_T = -(1/v)(\partial v/\partial p)_T$ and $\alpha_p = (1/v)(\partial v/\partial T)_p$.

LECTURE X (STAT)

We have finished the thermodynamic part of this Course. Thermodynamics is a phenomenological theory the main aim of which is establishing relations between measurable quantities characterizing macroscopic properties of bodies (systems). It also formulates general rules allowig to tell which processes are possible and imposes limits on possible useful works that can be extracted from macroscopic systems. Results obtained within this theory are very general and are always true. The central role in the structure of thermodynamics is played by entropy which, however, does not find a natural interpretation within this theory itself. Furthermore, making quantitative predictions on the basis of thermodynamics requires that certain functions (thermodynamic potentials) be known, but within the formalism of pure thermodynamics these functions cannot be computed - they have to be laboriously reconstructed from experimental data (which, however, in applied and technical sciences is a good and fruitful method!).

We now turn to statistical physics (or statistical mechanics) the first (but not the only one!) goal of which is the derivation of thermodynamics or, putting it differently, providing methods for computing equilibrium properties of various physical systems on the basis of more fundamental laws[137] (theories) governing the behaviour of microscopic constituents (molecules, atoms or physical fields) of these system. In this approach entropy will acquire a concrete interpretation of a measure of disorder or rather of a measure of the lack of information on the actual microscopic state of the system. Our primary task will therefore be understanding how this lack of information neverheless does not prevent making definite predictions concerning macroscopic quantities characterizing a given system and to provide concrete recipies for obtaining with the help of statistical methods thermodynamical potentials such as entropy, Helmholtz free energy etc. (depending on the conditions in which the system is placed) as functions of their natural variables on the basis of concrete microscopic theories.

Within the statistical approach it is, of course, possible to go much further and relying on microscopic dynamical theories investigate fluctuations of various quantities as well as their correlations. Statistical approach makes it also possible to consider systems not in equilibrium, in particular to study how equilibrium is reached in various situations. This part of the statistical method is called kinetic theory. In fact the whole statistical mechanics (also the equilibrium one) partly grew out from the Boltzmann and Maxwell studies of such problems within the kinetic theory of gases based on classical mechanics. The kinetic approach, essentially still based on the classical Boltzmann equation supplemented only with reaction rates computed within a suitable quantum theory, finds at present numerous applications for example in modern cosmology.

In this Course we will only discuss fundamentals of the equilibrium statistical thermordynamics (the one based on classical as well the one based on quantum mechanics).

---

[137] It should be clear from the beginning that the ultimate laws of Nature are not known yet (will they be ever known?), but they are largely irrelevant for understanding the behaviour of matter from the statistical standpoint: all that is needed is an effective theory sufficiently accurately predicting the behaviour of atoms and molecules.

Moreover, we will mainly treat only the simplest systems of mutually nointeracting elements (molecules, spins, etc.). Going further requires introducing special mathematical tools and in fact quickly leads to developing methods which are most easily understood only for those who have been trained on quantum field theory.

Thermodynamics was based on the assumption, which in fact is an idealization, but a one which macroscopically seems possible to be approached as closely as it is required, that by suitable walls every physical system can be isolated from any influences from without. A system isolated in this way attains then, after some time (which barring some exceptional systems is usually not too long on the macroscopic scale), an equilibrium state. Thermodynamics applies strictly speaking only to such equilibrium states. In contrast to the assumption on which thermodynamics is based, the most realistic assumption on which to base statistical mechanics is the statement that at the microscopic level no real system can be absolutely isolated with respect to random (uncontrolable) minute external perturbations. This is particularily true of that macroscopic part of a (classical) system on which measurements are being made and which for short periods (when the mesurements are made) can be treated as macroscopically isolated from the rest of the system. More fundamentally, one should adopt the view that, for reasons which will be given in the next Lecture, this is always true for any macroscopically large quantum system, so it is the right basis on which to base quantum statistical mechanics; its classical counterpart must be constructed in such a way as to be the limit of quantum statistical mechanics and (from the pragmatic point of view) its role is only to simplify calculational aspects of the statistical treatement of physical systems.

When one is interested (measures) in those few properties of a macroscopic system which characterize its equilibrium state, then the same their values could be obtained by measuring them on copies of the system which microscopically are in states different than is our system. In other words, typically the same macrostate of a given system is realized by a very large number of its different microstates and this is true of equilibrium macrostates (which we consider in this Course) as well as of its nonequilibrium macrostates. The notion of the microstate depends on whether at the microscopic level the system is treated classically or quantum-mechanically and will acquire proper meanings in both these cases in due course. The crucial consequence of the impossibility of putting the system in absolute (from the microscopic point of view) isolation is that its microstate can never be known sufficiently precisely: even if the system started in a well defined microstate it would immediately make innumerable transitions to different microstates as a result of being perturbed by randomly acting external agents. Thus the most reasonable physical assumption on which to base the foundations of statistical mechanics of a macroscopic physical system the energy and the deformative parameters of which are fixed (directly controlled) is, as will be argued, to assume that all its possible microstates which realize the same its macrostate are equally probable. This is called the equal probabilities postulate which can also be given a justification on the basis of the information theory.

Foundations of classical statistical mechanics

Microscopic constituents of matter - molecules, atoms, fields - obey the laws of quantum mechanics.[138] For this reason it is certainly well justified to claim that the foundations of equilibrium and nonequilibrium statistical physics should be based only on this theory. While this is principially true, it is nevertheless worthwhile to consider statistical physics based on classical mechanics. Although some ingredients neessary to fully acccount for the behaviour of real physical systems are lost in this way (lost for example are spin degrees of freedom; indistinguishability of identical particles can only approximately be accounted for, phenomena at very low, $T \to 0$, temperatures cannot be captured correctly, etc.), developing statistical approach based on classical mechanics is not completely useless. Firstly, as it is closer to our intuition, it allows to better understand the basic principles and rules. Secondly, it is also useful in practice: for instance, it turns out that if properties of gases at not too low temperatures are considered, the contributions to the thermodynamic potentials of the motions of centers of masses of the molecules can always be treated classically (from the quantum mechanics point of view this motion is quasi-classical) and combined with the contributions of the internal motions of these molecules (rotations and vibrations) which (at least in some ranges of the temperature) must be treated quantum mechanically.

Thus we begin by considering a system consisting of a fixed number N of particles (molecules) which obey the laws of classical mechanics. A microstatate of such a system is in principle fully specified by giving instantaneous values of $3N$ generalized coordinates q^i and $3N$ velocities $\dot{q}^i$ (together they constitute the complete set of initial data for the Euler-Lagrange equations which would uniquely determine further evolution of the system if it were completely isolated). However, as the classical statistical mechanics should eventually emerge as a limiting case of quantum statistical mechanics which necessarily must be formulated in terms of Hamiltonians and the so-called canonical variables, it is much more convenient to specify the classical system's microstate by giving instantaneous values of $3N$ coordinates q^i and $3N$ conjugated momenta p_i and to adopt the Hamilton's equations as the ones which would determine the time evolution of an isolated system. The actual state of a system of N particles will be therefore represented by a point in a $6N$ dimensional **phase space** $\Gamma_{(q,p)}$ the $3N$ axes of which represent the system's positions q^i and the remaining $3N$ axes represent its momenta p_i. As these change with time, the point representing the system moves through the phase space $\Gamma_{(q,p)}$ (irrespectively of whether the system is microscopically isolated or not).

A macroscopic quantity[139] (an observable) which characterizes a given system depends, of course, on its actual microscopic state and can, therefore, be represented by a function $O = O(q,p)$ of its generalized coordinates and momenta. Because of some inertia of all macroscopic devices, experimental measurments made on the system do not really

---

[138] Or, at a yet deeper level penetrated (experimentally) to date, by the laws of quantum field theory which is in a sense (it takes me rougly half of the semester of teaching to explain this properly) the application of the general principles quantum mechanics to systems of relativistic fields or to systems of relativistic particles.

[139] We have in mind here a thermodynamic quantity of a mechanical nature, like pressure.

measure instantaneous values of the observable O but rather its time average of the sort

$$O_{\rm obs} = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} dt\, O(q,p)\,, \tag{210}$$

in which τ is sufficiently long on the microscopic scale to smooth out fluctuations. Obviously, the formula (210) stays valid whether the system is completely isolated or not. If the real system were completely isolated, the time evolution of its coordinates q^i and momenta p_i would be determined by the canonical Hamilton's equations of motion

$$\dot{q}^i = \frac{\partial\mathscr{H}(q,p)}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial\mathscr{H}(q,p)}{\partial q^i}\,, \qquad i = 1,\dots,3N, \tag{211}$$

and the initial conditions $q^i(t_0)$, $p_i(t_0)$ set at some instant t_0. The measured quantity (210) could be then in principle computed in terms of these initial data as

$$O = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} dt\, O(q(t),\, p(t))\,. \tag{212}$$

It is clear, however, that in the case of a real system this is not feasible: firstly, although owing to the growing power of computers one is at present able to simulate numerically the classical time evolution of systems consisting of, say, $N \sim 10^4$ particles (such simulations are done in modeling e.g the formation of structures - galaxies, clusters of galaxies, etc. - in the expanding Universe), simulating the time evolution of systems consisting of $N \sim N_A = 6.022\times10^{23}$ particles is certainly beoynd reach. Secondly, and this ultimately turns out to be decisive, one can never know the initial values $q^i(t_0)$, $p_i(t_0)$ of the coordinates q^i and momenta p_i.

Even worse is the situation when the system is not isolated from the microscopic point of view (or even not isolated macroscopically, when e.g. being in equilibrium it remains in thermal contact with a heat bath). In this case the time evolution of the variables q^i and p_i is not determined by a Hamiltonian depending solely only on these variables - to determine the evolution of the variables q^i and p_i of the considered system one would have to follow also the evolution of the variables specifying the state of the entire surrounding.

In view of this situation one has to resort to the statistical method (suggested by Gibbs) which employs **statistical ensembles**. By the ensemble one means a set of a very large number $\mathcal{N}$ (implicitly the limit $\mathcal{N}\to\infty$ is to be taken) of copies of **microscopically isolated** systems which are all macroscopically indistinguishable from the considered real system (which itself may be not isolated microscopically or even macroscopically, when it is in contact with e.g. a heat bath, or may even not be in equilibrium) at an initial instant of time - that is, all have their macroscopic characteristics (these are rather easy to specify if the considered real system is in full equlibrium as a macroscopically isolated system, or in equilibrium with its surrounding; considerably more difficult is the specification of macroscopic characteristics of a system which is not in equilibrium) identical with those of the considered real system at some $t_{\rm in}$. For example, the ensemble representing at the moment $t_{\rm in}$ a real system in equlibrium (either a macroscopically isolated system or in equilibrium with its surrounding) the macrostate of which does not change in time, could be constructed by taking all points of the phase space $\Gamma_{(q,p)}$ which this system "visits" in consequtive instants $t_{\rm in}+n\Delta t$ with some Δt and $n\to\infty$ and assigning them all as the initial data at the instant $t_{\rm in}$ to members of the ensemble. More generally (not restricting to systems in equilibrium), one could operationally construct an ensemble by preparing $\mathcal{N}$ copies of identical systems in the same way as the system itself (in the same conditions) and putting all of them in perfect isolation at some initial moment $t_{\rm in}$. The systems of the ensemble constucted according to these prescriptions would populate in the phase space $\Gamma_{(q,p)}$ different possible microstates which give rise to the same macrostate (as specified by its macroscopically measured characteristics) and can be taken to represent the real system. However in view of the fact that the construction of the ensemble along the indicated lines evidently cannot be carried out in practice, the task central to the statistical method is to **theoretically construct** an ensemble (that is, in statistical physics bases on classical mechanics, to select a distribution function $\rho(q,p,t_{\rm in})$ - see below) in such a way that the distribution of its systems over the phase space $\Gamma_{(q,p)}$ (at the instant $t_{\rm in}$) is, from the macroscopic point of view, representative for the considered real system (at the same instant $t_{\rm in}$). It will be seen that in the case of systems in equilibrium there exist definite rules of achieving this.

In general ensembles obtained either operationally in the way specified above or constructed theoretically can be characterized by their **phase space distribution function** $\rho(q,p,t)$ normalized to unity[140]

$$\int_\Gamma d\Gamma_{(q,p)}\,\rho(q,p,t)=1\,,\qquad(213)$$

(the precise form - its normalization factor - of the measure $d\Gamma_{(q,p)}\propto d^{3N}q\,d^{3N}p$ will be specified later). It gives the fraction of the number of the ensemble systems which at the instant t are contained in the infinitesimal phase space volume element $d\Gamma_{(q,p)}$ around the point (q^i,p_i). In other words,

$$\mathcal{N}\rho(q,p,t)\,d\Gamma_{(q,p)}\,,$$

is the number of systems of the ensemble contained (at the instant t) in this infinitesimal phase space volume element. Given the distribution function $\rho(q,p,t)$, with every observable quantity O which is determined by the microscopic state of the system associated can be the ensemble average $\overline{O}(t)$ over the phase space:

$$\overline{O}(t)\equiv\int_\Gamma d\Gamma_{(q,p)}\,\rho(q,p,t)\,O(q,p)\,.\qquad(214)$$

It is the ensemble average which in the statistical approach is going to replace the time averages (210) and the justification of this replacement will be discussed (to the extent to which it is possible) below.

---

[140] We restrict ourselves here to real systems not exchanging matter with their surrounding. Extension to open systems (in equlibrium) will be made later.

Since the systems forming an ensemble are by definition constructed as (microscopically) isolated, the time evolution of each of them can be analyzed with the help of the canonical equations of motion (211) with the Hamiltonian $\mathscr{H}(q,p)$ which includes only the internal interactions of the system (the terms $\mathscr{H}_{\rm surr}$ and $\mathscr{H}_{\rm syst-surr}$ are absent from it) and defines therefore the dynamics of an autonomous system. As each of the systems of the ensemble evolves in time, also, in general, changes the distribution function ρ globally characterizing their distribution over the phase space. Certain global features of these changes of ρ can be understood on the basis of the Liouville theorem.

Liouville theorem and its consequences
The Liouville theorem is a technical result obtained within classical mechanics[141] which provides a global characterization of the dynamical behaviour of a large number of identical systems forming an ensemble. It is needed in the development of the argument justifying the statistical approach.

Let us consider first the total differential of the introduced distribution function:

$$d\rho = \frac{\partial\rho}{\partial t}\,dt + \sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial q^i}\,dq^i + \frac{\partial\rho}{\partial p_i}\,dp_i\right).$$

The sum $\rho + d\rho$ is obviously the value of the distribution function at the neighbouring point $(q^i + dq^i,\, p_i + dp_i)$ of an arbitrarily chosen point $(q^i,\, p_i)$ of the phase space at the moment $t + dt$, written in terms of its value ρ at the point $(q^i,\, p_i)$ at the instant t. One can, however, ask what will be the value of the distribution function at the instant $t + dt$ at the point $(q^i + dq^i,\, p_i + dp_i)$ which is related to $(q^i,\, p_i)$ by the motion of the system. This requires correlating dq^i and dp_i with dt by $dq^i = \dot{q}^i dt$, $dp_i = \dot{p}_i dt$. The distribution function at this neighbouring point linked to $(q^i,\, p_i)$ by the dynamics is equal $\rho + (d\rho/dt)dt$ with (the second form follows by using the canonical equations (211) and the definition of the Poisson bracket denoted "PB")

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial q^i}\,\dot{q}^i + \frac{\partial\rho}{\partial p_i}\,\dot{p}_i\right) \equiv \frac{\partial\rho}{\partial t} + \{\rho,\, \mathscr{H}\}_{\rm PB}\,. \tag{215}$$

The first part of the Liouville theorem states that $d\rho/dt$ is zero.

To prove this we consider an arbitrarily chosen finite (i.e. of finite measure, in the mathematical language) domain $\Delta\Gamma_{(q,p)}$ of the phase space $\Gamma_{(q,p)}$ and ask how the instantaneous number

$$\mathcal{N}_{\Delta\Gamma}(t) = \int_{\Delta\Gamma} d\Gamma_{(q,p)}\,\mathcal{N}\rho(q,p,t)\,,$$

of phase points (representing the systems of the ensemble) contained in this domain changes between t and $t + dt$. On one hand, this change is simply given by $(d\mathcal{N}_{\Delta\Gamma}(t)/dt)dt$

---

[141] It has, of course, also a counterpart in quantum mechanics.

where

$$\frac{d\mathcal{N}_{\Delta\Gamma}(t)}{dt} = \frac{d}{dt}\int_{\Delta\Gamma} d\Gamma_{(q,p)}\,\mathcal{N}\rho(q,p,t) = \int_{\Delta\Gamma} d\Gamma_{(q,p)}\,\mathcal{N}\,\frac{\partial\rho(q,p,t)}{\partial t}\,.$$

On the other hand, the same change $d\mathcal{N}_{\Delta\Gamma}(t)$ must be equal to the flux of the phase points, given by the $6N$ dimensional flux "vector"

$$\mathcal{N}\rho\,\mathbf{u} \equiv \mathcal{N}\rho\begin{pmatrix}\dot{q}^i\\ \dot{p}_i\end{pmatrix},$$

(which is the product of $\mathcal{N}\rho$ itself and the $6N$ dimensional "velocity vector" $\mathbf{u}$), integrated over the boundary of the domain $\Delta\Gamma_{(q,p)}$ and multiplied by dt ($d\mathbf{\Sigma}$ is the $6N-1$-dimensional differential area vector normal to the surface and directed outwards):

$$d\mathcal{N}_{\Delta\Gamma}(t) = -dt\int_{\partial\Delta\Gamma} d\mathbf{\Sigma}\cdot\mathbf{u}\,\mathcal{N}\rho\,.$$

By the Stokes theorem this can be written in the form ($\boldsymbol{\nabla}$ stands for the $6N$-dimensional differential operator)

$$\frac{dN_{\Delta\Gamma}(t)}{dt} = -\int_{\Delta\Gamma} d\Gamma_{(q,p)}\boldsymbol{\nabla}\cdot(\mathbf{u}\,\mathcal{N}\rho)\,.$$

Equating the two forms of the rate $d\mathcal{N}_{\Delta\Gamma}(t)/dt$ and taking into account that the domain $\Delta\Gamma_{(q,p)}$ was arbitrary (which implies that the integral equality can be replaced by its local differential version), we obtain the equality

$$\begin{aligned}0 = \frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot(\mathbf{u}\,\rho) &= \frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left[\frac{\partial}{\partial q^i}\,(\dot{q}^i\rho) + \frac{\partial}{\partial p_i}\,(\dot{p}_i\rho)\right]\\ &= \frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left[\frac{\partial\rho}{\partial q^i}\,\dot{q}^i + \frac{\partial\rho}{\partial p_i}\,\dot{p}_i + \left(\frac{\partial\dot{q}^i}{\partial q^i} + \frac{\partial\dot{p}_i}{\partial p_i}\right)\rho\right].\end{aligned}$$

Since

$$\frac{\partial\dot{q}^i}{\partial q^i} = \frac{\partial}{\partial q^i}\,\frac{\partial\mathscr{H}}{\partial p_i} = \frac{\partial^2\mathscr{H}}{\partial q^i\partial p_i}\,,\qquad \frac{\partial\dot{p}_i}{\partial p_i} = -\frac{\partial}{\partial p_i}\,\frac{\partial\mathscr{H}}{\partial q^i} = -\frac{\partial^2\mathscr{H}}{\partial p_i\partial q^i}\,,$$

the last term in the square brackets is zero and, comparing with (215), we indeed conclude that $d\rho/dt = 0$. The vanishing of the total time derivative of the distribution function ρ means that this function is constant along trajectories of the phase points or, that the "phase fluid is incompressible". Alternatively, this can be written as the equality

$$\rho(q,p,t) = \rho(q_0,p_0,t_0)\,, \tag{216}$$

holding for (q^i,p_i) related at $t = t_0 + s$ to the values (q^i_0,p^0_i) of the phase space coordinates at t_0 by the motion: $q^i = q^i(s,q_0,p_0,t_0)$ and $p_i = p_i(s,q_0,p_0,t_0)$.

To obtain the second part of the Liouville theorem we consider a small (infinitesimal in fact) domain $\Delta\Gamma_{(q,p)}$ of the phase space and follow over the interval s of time the motion of all the phase points which were contained in it at some instant t_0. At the moment $t = t_0 + s$ these points will occupy another domain of the phase space. The shape of this new domain is marked by the points which at t_0 formed the boundary of $\Delta\Gamma_{(q,p)}$. Moreover, no one point of those which at t_0 formed the interior of $\Delta\Gamma_{(q,p)}$ could have evolved outside of the new domain, because in doing so it would have to cross on the way the moving boundary of the domain: at some moment between t_0 and $t_0 + s$ it would have to coincide with some boundary point; but since the Hamilton's equations of motion are causal and their solutions are unique, it would have to coincide at every other instant with this point of the boundary. It follows that the number of phase points inside the migrating domain could not change and since we have shown that along the phase trajectories the distribution function (i.e. the density of the points in the phase space) is constant, we conclude that although the shape of the moving domain could have (even drastically) changed between t_0 and $t_0 + s$, its volume remained unchanged. This in turn means that

$$J = \frac{\partial(q,p)}{\partial(q_0,p_0)} = 1\,. \tag{217}$$

where J is the jacobian of the change (transformation) of the integration variables from (q^i, p_i) to (q^i_0, p^0_i) which is given by the motion: $q^i = q^i(s, q_0, p_0)$, $p_i = p_i(s, q_0, p_0)$ (with the time s playing the role of an arbitrary transformation parameter).

Put differently, the argument given above shows that

$$\int_{\Delta\Gamma(t_0+s)} d\Gamma_{(q,p)}\,\mathcal{N}\rho(q,p,t_0+s) = \int_{\Delta\Gamma(t_0)} d\Gamma_{(q_0,p_0)}\,\mathcal{N}\rho(q_0,p_0,t_0)\,. \tag{218}$$

After using the formulae: $q^i = q^i(s, q_0, p_0, t_0)$, $p_i = p_i(s, q_0, p_0, t_0)$ following from the system's time evolution, as defining the transformation of the integration variables and using the equality (216) to replace $\rho(q, p, t_0 + s)$ with $\rho(q_0, p_0, t_0)$, the left integral takes the form

$$\int_{\Delta\Gamma(t_0)} d\Gamma_{(q_0,p_0)}\,\frac{\partial(q,p)}{\partial(q_0,p_0)}\,\mathcal{N}\rho(q_0,p_0,t_0)\,,$$

Its equality to the integral on the right side of (218) means therefore that (217) holds.[142]

Correspondence of the time and ensemble averages

We now make the crucial step in developing the statistical approach. We are going to postulate the relation between the quantity (210)

$$O_{\text{obs}} = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} dt\,O(q,p)\,,$$

---

[142] The same follows also from the general fact that the transformation $q^i = q^i(s, q_0, p_0, t_0)$, $p_i = p_i(s, q_0, p_0, t_0)$ from the set (q^i_0, p^0_i) of the canonical variables to the new variables (q^i, p_i) defined by the actual motion of the system is a canonical transformation and as such must have unit Jacobian.

measured on the real system (isolated macroscopically or not) and the appropriate ensemble average. The latter is defined as follows. We construct an ensemble which at $t_{\rm in}$ is representative for the considered real system. With each (isolated) system of the ensemble which from its microstate (q^i_0, p^0_i) at $t_0 \geq t_{\rm in}$ evolves as dictate the autonomous Hamilton's equations (211):

$$q^i = q^i(s, q_0, p_0)\,, \qquad p_i = p_i(s, q_0, p_0)\,,$$

the time average

$$O_\tau(q_0, p_0) = \frac{1}{\tau}\int_0^\tau ds\, O(q(s, q_0, p_0),\, p(s, q_0, p_0))\,, \tag{219}$$

can be associated. The period τ (here and in (210)) can be taken sufficiently long from the microscopic point of view (for instance, in the case of a gas τ should be large compared to the mean time between collisions), so that O_τ is practically independent of τ (though, as indicated by the notation, it can still depend on the initial point (q^i_0, p^0_i)). If the system is not in equlibrium, τ cannot be too long too, if the formulae are to represent a measurement which pertains to a definite stage of the system's evolution (towards an equilibrium state). The main postulate is then that $O_{\rm obs}$ should be identified with the ensemble average

$$\overline{O_\tau} \equiv \int d\Gamma_{(q_0,p_0)}\, \rho(q_0, p_0, t_0)\, O_\tau(q_0, p_0)\,. \tag{220}$$

The validity of the identification of $O_{\rm obs}$ with $\overline{O_\tau}$ depends on the magnitude of the fluctuations of the values of O_τ's (computed starting from different (q_0, p_0)'s) around the mean $\overline{O_\tau}$. If the real system is completely (microscopically) isolated (which is the limiting, not very realistic, case), it can be taken as one of the systems of the ensemble. Smallness of the fluctuations means that the probability that $O_\tau(q_0, p_0)$ computed taking a randomly chosen element of the ensemble - in particular taking the real system - markedly deviates from $\overline{O_\tau}$ is very small, so the same must be true of $O_{\rm obs}$. If the real system is not isolated (either only microscopically or because it is in contact with its surrounding), then its phase space trajectory can be imagined to be piecewise composed out of parts of trajectories traced out by different systems (which are microscopically isolated) of the ensemble - due to random external perturbations the real system jumps from time to time from a phase space trajectory of one isolated system onto the trajectory of another isolated system. It again follows (or, better, can be maintained) that the probability that $O_{\rm obs}$ which is a composition (in varying proportions) of parts of different O_τ's deviates significantly from the ensemble average $\overline{O_\tau}$ is small if the fluctuations (which are determined by the ensemble) of O_τ's given by (219) are small provided the ensemble represents the real system sufficiently faithfully.[143] In either case, the validity of the identification of $O_{\rm obs}$

---

[143] This is why the problem of choosing the representative ensemble is the central one of the whole statistical approach. Fortunately simple choices which are made in typical situations of macroscopically isolated systems, or systems exchanging energy with a heat bath prove (by the results they lead to) to be sufficiently adequate.

with the ensemble average $\overline{O_\tau}$ is reduced to the problem whether the fluctuations of O_τ's around $\overline{O_\tau}$ are small, which can be investigated entirely in the framework of the ensemble itself, without any further reference to the real system. It is expected that typically the fluctuations are indeed small owing to the smoothing effect of time averaging involved in computing individual O_τ's and owing to the enormous number of degrees of freedom involved.

The arguments given above are by no means rigorous. It is clear that the value of $\overline{O_\tau}$ depends on the form of the distribution function $\rho(q,p,t)$ which, at t_0 must reasonably well represent the real system. It can be however expected that the mean values $\overline{O_\tau}$ of observable quantities O (but not their fluctuations) are to some extent insensitive to the form of ρ. Indeed, one knows that e.g. a system in equilibrium in contact with a heat bath at a temperature T and an identical system in complete isolation but having the same temperature T are thermodynamically equivalent although the distribution functions associated with them will be different.

The construction of the statistical ensembles and the prescription for computing the ensemble average $\overline{O_\tau}$, which according to the adopted postulate is to be identified with the really measured value $O_{\rm obs}$ given above, are fairly general and apply to (real) systems in equlibrium (either as macroscopically isolated system or with their surrounding) as well as to systems out of equilibrium. If, as in this Course, we are interested only in systems in equilibrium, some simplifications can be made. Firstly, the measured quantities $O_{\rm obs}$ are in this case independent of time (macroscopic characteristics of systems in equilibrium by definition do not vary in time). Therefore also the ensemble averages $\overline{O_\tau}$ given by (220) should be independent of t_0. It is clear that this requires that the corresponding ensemble distribution function $\rho(q,p,t)$ does not depend explicitly on time:

$$\frac{\partial}{\partial t}\,\rho(q,p,t)=0\,. \tag{221}$$

Since the Liuoville theorem states that $d\rho/dt=0$, from the formula (215) it follows that the distribution functions ρ of the ensembles intended to represent systems in equilibrium, in addition to being not explicitly dependent on time, must have zero Poisson bracket with the Hamiltonian of the system (treated as isolated):

$$\{\rho(q,p),\,\mathscr{H}(q,p)\}_{\rm PB}=0\,. \tag{222}$$

This means that the distribution function can depend on the dynamical variables q^i and p_i only through conserved quantities specific for the considered system. Most of complex (isolated) systems consisting of very large numbers of particles have only a few of conserved quantities:[144] energy, represented by the Hamiltonian itself, and perhaps the total momentum and total angular momentum of which the last two are usually eliminated by enclosing the system in suitable spatial walls. Thus typically the distribution function of

---

[144] The exception are so called integrable systems which can have up to $6N$ (completely integrable systems) conserved quantities. We will not be concerned here with such systems.

an ensemble representing a real system in equilibrium has the form

$$\rho(q,p,t)=\rho(q,p)=\rho(\mathscr{H}(q,p))\,. \tag{223}$$

Moreover, the recipe for calculating the ensemble average (220) pertaining to the properties of a system in equilibrium can be considerably simplified: in the explicit formula

$$\begin{aligned}\overline{O_\tau}&=\frac{1}{\tau}\int d\Gamma_{(q_0,p_0)}\,\rho(q_0,p_0)\int_0^\tau ds\,O(q(s,q_0,p_0),\,p(s,q_0,p_0))\\&=\frac{1}{\tau}\int_0^\tau ds\int d\Gamma_{(q_0,p_0)}\,\rho(q_0,p_0)\,O(q(s,q_0,p_0),\,p(s,q_0,p_0))\,,\end{aligned}$$

one can use the result (216) with $q^i=q^i(s,q_0,p_0)$, $p_i=p_i(s,q_0,p_0)$ specified to distribution fuctions satisfying (221) and the result (217) to replace $d\Gamma_{(q_0,p_0)}$ by $d\Gamma_{(q,p)}$. After these operations the integral over $d\tau$ factorizes completely, cancels out against the explicit $1/\tau$ factor and one obtains an explicitly τ-independent formula

$$\overline{O_\tau}\equiv\overline{O}=\int d\Gamma_{(q,p)}\,\rho(q,p)\,O(q,p)\,. \tag{224}$$

In the case of systems in equilibrium, when taking the limit $\tau\to\infty$ in the definition (210) of the measured quantity is allowed (because the measured macroscopic properties do not depend in this case on time) the formula (224) can be also justified à la Landau: in this case one can imagine that the phase space trajectory of the real system is followed over a period τ and the time intervals $\Delta t_{(q,p)}(\tau)$ the trajectory spends in different small domains $\Delta\Gamma_{(q,p)}$ of the phase space are counted. If the equilibrium distribution function is constructed by taking

$$\rho(q,p)\propto\lim_{\tau\to\infty}\frac{\Delta t_{(q,p)}(\tau)}{\tau}\,, \tag{225}$$

and normalizing the resulting function of q^i and p_i to unity with respect to the measure $d\Gamma_{(q,p)}$, the equality

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau dt\,O(q(t),\,p(t))=\int d\Gamma_{(q,p)}\,\rho(q,p)\,O(q,p)\,,$$

follows automatically because the left integral over dt can be (somewhat heuristically) written as

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau dt\,O(q(t),\,p(t))=\lim_{\tau\to\infty}\sum_{(q,p)}\frac{\Delta t_{(q,p)}}{\tau}\,O(q,p)\,,$$

with the sum on the right hand side taken over cells of the phase space in which the system spends overall periods $\Delta t_{(q,p)}$ in its journey over the phase space. Of course also in this derivation the question how to theoretically construct the distribution ρ similar to the one defined here operationally is not solved by the above reasoning.

Once the distribution function $\rho(q,p)$ of the ensemble representing a real system in equilibrium is given, obtaining theoretical predictions for measured macroscopic quantities characterizing this system reduces to the application of the probability theory: the phase space parametrized by the variables (q^i, p_i) plays the role of the space of elementary events, $\rho(q,p)d\Gamma_{(q,p)}$ plays the role of the probability distribution and observables $O(q,p)$ become random variables defined on the space of elementary events. The probability distribution $\rho_O(O)$ of a random variable (observable) O can be then constructed according to the standard prescription[145]

$$\rho_O(O) = \int d\Gamma_{(q,p)}\, \rho(q,p)\, \delta(O(q,p) - O)\,. \tag{226}$$

Owing to the properties of the Dirac delta function, the probability distribution $\rho_O(O)$ constructed in this way is automatically normalized to unity. The mean value of O can be then computed either as in (224) or, as

$$\overline{O} = \int dO\, O\, \rho_O(O)\,. \tag{227}$$

The standard estimate of the fluctuations of the random variable O around its mean value $\overline{O}$ is obtained by computing the quantity

$$\sigma_O^2 \equiv \overline{(O - \overline{O})^2} = \int dO\, (O - \overline{O})^2\, \rho_O(O) = \overline{O^2} - \overline{O}^2\,. \tag{228}$$

called the **mean quadratic fluctuation** and the taking its square root. The (dimensionless) measure of fluctuations is the relative fluctuation given by the ratio

$$\frac{\sqrt{\sigma_O^2}}{\overline{O}}\,, \tag{229}$$

which tells how large fluctuations are compared to the mean value.[146] The success of the statistical method applied to large ($N \sim N_A$) systems relies mostly on the fact that even if the distribution $\rho(q,p)$ itself is flat (as in the case of the microcanonical ensemble - to be defined below), the probability distribution $\rho_O(O)$ of any macroscopic observable has in most cases[147] an enormously sharp peak at a value O^* which is therefore almost the same as $\overline{O}$ and the relative fluctuations are very tiny indeed.

---

[145] This prescription generalizes the ordinary summation of probabilities of discrete elementary events leading to the same value of the random variable; for instance, if the value of the random variable f is -1 for an odd number number of dots on a (perfectly symmetric) dice and and $f = +1$ for an even number, the probabilities of the values $f = \mp 1$ are $p(-1) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$ and $p(+1) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$. In the case of continuos probability distributions the integration over $d\Gamma_{(q,p)}$ combined with the Dirac delta function essentially sums up probabilities of elementary events (q^i, p_i) which lead to the same value of the observable O.

[146] Of course, if $\overline{O} = 0$, one has to invent another dimensionless measure of fluctuations.

[147] But not in all cases: for instance, the distribution function of the density of a fluid at the critical point does not have this property.

Statistical independence
If the system (the space of elementary events in the general probability theory) can be split into two mutually (almost) noninteracting parts a and b (two sets the elementary events belonging to them are mutually unrelated) which can be treated as not influencing one another in any significant way, these systems should be statistically independent: their joint distribution function $\rho_{a,b}(q_a, p_a, q_b, p_b, t)$ should factorize, that is, the following equality should hold

$$d\Gamma_{(q_a,p_a,q_b,p_b)}\, \rho_{a,b}(q_a, p_a, q_b, p_b, t) = d\Gamma_{(q_a,p_a)}\, \rho_a(q_a, p_a, t)\, d\Gamma_{(q_b,p_b)}\, \rho_b(q_b, p_b, t)\,. \tag{230}$$

The observables $O_a = O(q_a, p_a)$ and $O_b = O(q_b, p_b)$ pertaining to these two subsystems are then independent which for example means that

$$\overline{O_a O_b} = \overline{O_a}\; \overline{O_b}\,,$$

etc.

Statistical independence can be invoked to obtain two important results concerning macroscopic systems (bodies). Firstly, if the system is large it can be mentally divided into large macroscopic parts which are to a good approximation statistically independent.[148] It is natural to assume that the number of statistically independent parts (which can still be treated as macroscopic) into which the system is divided remains in a proportion to the number N of its constitutive elements (molecules, for instance). If one is interested in an additive (extensive) quantity O characterizing the system - for instance the system is in equilibrium with the heat bath and therefore one is interested in its internal energy, or when the system is in mechanical contact with the atmosphere and one asks about its total volume - then from the statistical independence of its macroscopic parts it immediately follows that the relative fluctuation (229) of this quantity decreases with the size of the body. Indeed, an extensive quantity O can be split into the sum

$$O = \sum_a O_a\,,$$

of quantities characterizing its macroscopic parts (labeled by a) and by the argument given above,

$$\overline{O} = \sum_a \overline{O_a} \propto N\,,$$

because all the parts are essentially identical. Moreover,

$$\sigma_O^2 = \overline{(O - \overline{O}\,)^2} = \overline{\left(\sum_a (O_a - \overline{O_a}\,)\right)^2} = \sum_a \sum_b \overline{(O_a - \overline{O_a}\,)(O_b - \overline{O_b}\,)}\,.$$

---

[148] Statistical independence holds in such a case over not too long periods; in the case of systems in equilibrium the macroscopic states of which do not change in time, it should however be always (approximately) true.

In the double sum one can now single out the terms in which $a = b$ and, appealing to the statistical independence, write

$$\begin{aligned}\sigma_O^2 &= \sum_a \overline{(O_a - \overline{O_a}\,)^2} + \sum_{a\neq b} \overline{(O_a - \overline{O_a}\,)(O_b - \overline{O_b}\,)} \\ &= \sum_a \overline{(O_a - \overline{O_a}\,)^2} + \sum_{a\neq b} \overline{(O_a - \overline{O_a}\,)}\ \overline{(O_b - \overline{O_b}\,)}\,. \end{aligned} \tag{231}$$

The second sum vanishes because $\overline{O_b - \overline{O_b}} = \overline{O_b} - \overline{O_b} = 0$ and the first sum, since all parts into which the system has been split are essentially identical, is proportional to N. It follows therefore, that the relative fluctuation of the extensive quantity scales as

$$\frac{\sqrt{\sigma_O^2}}{\overline{O}} \propto \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}\,, \tag{232}$$

and is really minute, if the system is large ($N \sim N_A$, where N_A is the Avogadro number).

Another important consequence of the statistical independence which is implied by the factorization (230) of the distribution function is as follows. Suppose two subsystems, a and b, of a large real system in an equilibrium state can be treated as independent because, as in the considerations above, they are two macroscopic parts of the same macroscopic body. The Hamiltonian of the members of the ensemble (which, recall it once more, are completely isolated systems) which represents the entire real system in equilibrium is then just the sum $\mathscr{H} = \mathscr{H}_a + \mathscr{H}_b$ ($\mathscr{H}_a$ depends on the variables of the subsystem a, while $\mathscr{H}_b$ on the variables of the subsystem b). In view of the fact that the subsystems a and b of the members of the ensemble should in this case be statistically independent, the distribution function of the ensemble representing the entire system must have the property (recall the result (223)!)

$$\rho_{a,b}(\mathscr{H}) = \rho_{a,b}(\mathscr{H}_a + \mathscr{H}_b) = \rho_a(\mathscr{H}_a)\,\rho_b(\mathscr{H}_b)\,,$$

or

$$\ln\rho_{a,b}(\mathscr{H}_a + \mathscr{H}_b) = \ln\rho_a(\mathscr{H}_a) + \ln\rho_b(\mathscr{H}_b)\,. \tag{233}$$

The only possibility of satisfying this requirement is[149]

$$\ln\rho_a(\mathscr{H}_a) = \alpha_a - \beta\mathscr{H}_a\,, \tag{234}$$

or, in the case the number of molecules becomes a dynamical variable (when the real system is in equilibrium exchanging with its surrounding both energy and matter),

$$\ln\rho_a(\mathscr{H}_a) = \alpha_a - \beta\mathscr{H}_a + \gamma N_a\,. \tag{235}$$

---

[149] Indeed, differentiating the equality $\ln f(x+y) = \ln f(x) + \ln f(y)$ with respect to x one obtains that $f'(x+y)/f(x+y) = f'(x)/f(x)$; in the same way one finds that $f'(x+y)/f(x+y) = f'(y)/f(y)$; hence $f'(y)/f(y) = f'(x)/f(x) = -\beta = \text{const.}$, because both sides are functions of a different independent variable. From this the conclusion follows.

Thus, statistical independence of systems which can be treated as isolated from one another imposes stringent constraints on the possible forms of the distribution functions of statistical ensembles representing systems in equilibrium.

Classical Microcanonical Ensemble

We now introduce the distribution function of the statistical ensemble which should adequately represent the system which from the macroscopic point of view is isolated and in equilibrium or more precisely, is (macroscopically) adiathermally isolated and its deformative parameters have fixed values (are directly controlled). Because according to the point of view adopted here such a system is not isolated at the microscopic level, but is constantly perturbed by its surrounding, it is reasonable to assume that its internal energy is fixed only up to some uncertainty, which we will represent by allowing the system's energy to be in the range $(E, E+\Delta E)$ with $\Delta E \ll E$, and that all system's microstates with the energy in this range are **equally probable**. In agreement with this assumption the distribution function of the statistical ensemble representing a macroscopically isolated system is taken in the form

$$\rho(q,p) = \begin{cases} \text{Const.} & \text{if} \quad E \leq \mathscr{H}(q,p) \leq E+\Delta E \\ 0 & \text{otherwise} \end{cases} \tag{236}$$

The ensemble defined by this distribution function is called **microcanonical**. Since the distribution function ρ must be normalized to unity, the constant is equal $1/\Gamma(E,V,N,\Delta E)$ where

$$\Gamma(E,V,N,\Delta E) = \int_{E\leq\mathscr{H}(q,p)\leq E+\Delta E} d\Gamma_{(q,p)}\,, \tag{237}$$

is the volume of the corresponding part of the phase space. In various considerations useful are also two other quantities:

$$\Sigma(E,V,N) = \int_{\mathscr{H}(q,p)\leq E} d\Gamma_{(q,p)} \equiv \int d\Gamma_{(q,p)}\,\theta(E-\mathscr{H}(q,p))\,, \tag{238}$$

which gives the phase space volume corresponding to the system's total energy less than E and:

$$\omega(E,V,N) = \frac{\partial\Sigma(E,V,N)}{\partial E} = \int d\Gamma_{(q,p)}\,\delta(\mathscr{H}(q,p)-E)\,, \tag{239}$$

which gives the "area" of the shell corresponding to the system's total energy equal E; in the quantum case $\omega(E)$ will have the clear interpretation of the density of quantum states of the system. Since the energy allowance ΔE is small compared to E, a useful working approximation is

$$\begin{aligned}\Gamma(E,V,N,\Delta E) &= \Sigma(E,V,N) - \Sigma(E-\Delta E,V,N) \\ &\approx \frac{\partial\Sigma(E,V,N)}{\partial E}\,\Delta E = \omega(E,V,N)\,\Delta E\,.\end{aligned} \tag{240}$$

As the quantity $\Gamma(E, V, N, \Delta E)$ characterizing the spread of the systems of the microcanonical ensemble over the phase space is finite, one can define the **statistical entropy** by the simple formula[150]

$$S_{\rm stat} = k_{\rm B} \ln \Gamma(E, V, N, \Delta E)\,, \tag{241}$$

in which $k_{\rm B} = 8.617 \times 10^{-5}$ eV·K^{-1} is the Boltzmann constant. This generalizes to all systems by interpreting Γ as the "volume" of the phase space (the number of the microstates - see below) correspondng to the system's energy in the range $(E, E + \Delta E)$ and fixed values of the system's $o - 1$ deformative variables and possibly r different matter constituents:

$$S_{\rm stat} = k_{\rm B} \ln \Gamma(E, X_1, \dots, X_{o-1}, N_1, \dots, N_r, \Delta E)\,.$$

In this way the statistical entropy gives the measure of the "disorder" in the system, that is characterizes how much a given macrostate of the system is spread out over the systems' phase space, or how many microstates are associated with the given macrostate (specified by the total energy E, the volume V and the number N of particles). Of course it still remains to be shown how the statistical entropy $S_{\rm stat}$ defined by (241) is related to the thermodynamic entropy S, but the resonings needed for this follow closely those employed in showing that the Callenian entropy agrees with that introduced via the Clausius inequality. This will be done after we consider statistical mechanics based on quantum mechanics. The result will be that the thermodynamic entropy S should be identfied with

$$S = N \lim_{\infty} \left(\frac{S_{\rm stat}(E, V, N, \Delta E)}{N} \right), \tag{242}$$

where the symbol $\lim_{\infty}$ denotes the so-called **thermodynamic limit** which means $N \to \infty$, $E \to \infty$, $V \to \infty$ with the ratios E/N and V/N kept fixed. In this limit the dependence on ΔE drops out.[151] In most cases the statistical entropies computed replacing $\Gamma(E, V, N, \Delta E)$ under the logarithm in (241) by $\Sigma(E, V, N)$ or by $\omega(E, V, N)\Delta E$ lead in the thermodynamic limit to the same thermodynamic entropy.

Since the statistical entropy (241) is proportional to the logarithm of $\Gamma(E, V, N, \Delta E)$, it changes by an additive constant, when the normalization factor of the integration measure $d\Gamma_{(q,p)}$ is altered. In classical physics this factor is arbitrary. There is not such an ambiguity (apart from the choice of the energy allowance ΔE) in quantum statistical physics in which entropy (in the thermodynamic limit) is assigned unambigously to each

---

[150] This is the famous Boltzmann formula which in the form

$$S = k \ln W\,,$$

is engraved on his tomb auf dem Haupstad Wien Zentralfriedhof (at the Viennese cemetery).

[151] In classical statistical physics one could from the beginning set $\Delta E = 0$, but this is not so in quantum statistical mechanics (unless the system is peculiar). The dependence of S on ΔE drops out in the thermodynamic limit in both cases.

macrostate. It will be seen that if classical statistical mechanics is to be the limit of its quantum counterpart, the measure $d\Gamma_{(q,p)}$ appropriate for a system composed of N identical particle (molecules) should have the form

$$d\Gamma_{(q,p)} = \frac{d^{3N}q d^{3N}p}{N!\,(2\pi\hbar)^{3N}}\,. \tag{243}$$

The factor $(2\pi\hbar)^{3N}$ in the denominator makes the measure dimensionless. It is related to the rule (following e.g. from the old Bohr-Sommerfeld quantization condition) that in a one-dimensional quasi-classical motion of a single particle with every single quantum state associated can be the "area" $2\pi\hbar$ of the two-dimensional phase space. Thus $d^{3N}qd^{3N}p/(2\pi\hbar)^{3N}$ simply gives the number of $6N$ dimensional cells in the phase space available to the system each of which would correspond to one quantum state if the N particles were distinquishable. If particles are indistinguishable, the rules of quantum mechanics require that their quantum states be either symmetric (if particles are bosons - particles of integer spin) or antisymmetric (if they are fermions - particles of half-integer spin) with respect to interchanging particles (this will become more clear when we discuss the second quantization formalism). As a result the states which would be counted as distinct if the particles were distinguishable, are, when the particles are indistinguishable, one and the same quantum state. Therefore the number of quantum states available to particles is reduced roughly[152] by the factor of $N!$. This explains the origin of the extra factor of $N!$ in the denominator of the measure (243). This factor was first introduced by Gibbs (before its quantum origin became clear) to save extensiveness of entropy of the perfect gas which would not hold without it (this will be seen in classes; the problem is sometimes referred to as the Gibbs paradox). It should be also remarked that if the system consists of N_1 (indistinguishable) particles of one kind and of N_2 (indistinguishable) particles of another kind (the two kinds being different, so distinguishable) the appropriate measure is

$$d\Gamma_{(q,p)} = \frac{d^{3N_1}q_{(1)}d^{3N}p_{(1)}}{N_1!\,(2\pi\hbar)^{3N_1}}\,\frac{d^{3N_2}q_{(2)}d^{3N}p_{(2)}}{N_2!\,(2\pi\hbar)^{3N_2}}\,,$$

and it is precisely the factor $(N_1+N_2)!/N_1!N_2!$ by which it differs from the measure associated with the system of N_1+N_2 identical (indistinguishable) particles that leads to the entropy of mixing by which thermodynamic entropies of the two systems differ (Lecture VIII). With the normalization specified in (243) the Boltzmann formula (241) states that entropy (in units of $k_{\rm B}$) of a macroscopically isolated system is just the logarithm of the number of (microscopic) quantum states consistent with the macroscopic characteristics of (with the constraints imposed on) the system.

If the system is not macroscopically isolated and the distribution function of the representing it statistical ensemble is not localized in the phase space as is (236), the

---

[152] This does not allow to count the number of states absolutely correctly (it does not make distinction between systems of many bosons and fermions) but is a sufficiently good estimate in situations in which genuine quantum effects are not important for thermodynamics, that is, if the temperature is not too low.

entropy cannot be given by the formula (241) because the volume of the phase space avalilable to the system is infinite. One needs to invent then another measure of "disorder" in terms of which entropy can be defined. We will discuss this in general terms after we consider foundations of quantum statistical mechanics.

Ergodicity

It seems appropriate to close this Lecture by discussing briefly the so-called ergodic problem and the place it occupied in the past in the foundations of (classical) statistical mechanics. We have not made any reference to it in our considerations owing to our adoption of the point of view, which physically seems very natural and moreover allows to put the foundations of classical and quantum statistical mechanics (see the next Lecture) on the same basis, that no real system can be considered completely isolated at the microscopic level. It served us to justify (somewhat heuristically, let us agree!) the postulate of a priori equal probabilities assigned to all microstates accessible to a macroscopically isolated system (the energy and the deformative parameters of which are prescribed). This postulate can be introduced also without the necessity of relying (only) on the microscopic non-isolation of macroscopic systems by appealing to the information theory, if a somewhat more realistic view is taken on the possible outcomes of real measurements made on macroscopic systems. The microcanonical ensemble which results from this postulate leads to correct physical predictions and there can be no doubts about its correctness. However in the past many physicists subscribed to the position that a system which is isolated macroscopically is also microscopically strictly isolated against any influences from without, however minute (and the relevance of arguments based on the information theory were not yet commonly known), and were trying[153] to justify the postulate of a priori equal probabilities directly on the basis of microscopic time evolution of large autonomous isolated dynamical systems. It is in this context that the ergodicity problem appeared.

Briefly, the problem in which the notion of ergodicity has appeared was to show rigorously[154] that if the macroscopic system is isolated, the time average (210) with $\tau \to \infty$ is equivalent (that is, equal in this limit) to the ensemble average with the constant distribution function on the $6N-1$-dimensional hypersurface of constant energy. The proof of the equivalence has two parts. One is showing that a given point representing the time evolution of an isolated system in the phase space spends on its journey through the phase space in different domains $\Delta\Gamma_{(q,p)}$ the time intervals Δt which are proportional to the volumes of these domains.[155] The second part is showing that this is true independently of where in the phase space such a volume is located, that is that this proportionality holds (with the same proportionality constant) independently of which (initial) point the

---

[153] At least those who were not contented with the results the postulate leads to as its ultimate justification.

[154] And this is what attracted pure mathematicians to this field so that it effectively became a branch of mathematics.

[155] We will simplify slightly the problem and will consider the $6N$-dimensional domains assuming implicitly that the limit $\Delta E \to 0$ is taken only at the very end; otherwise we would have to consider the projections of the measure $d\Gamma_{(q,p)}$ onto the hypersurface of constant energy.

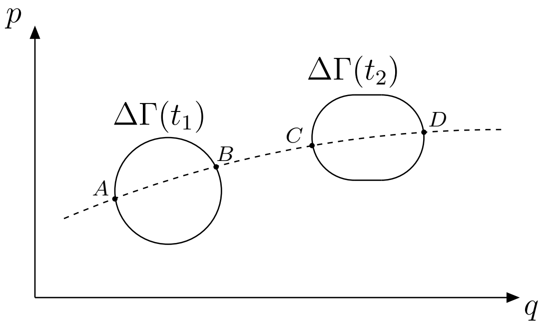

Figure 22: Migration over the phase space of a set of points.

trajectory is followed. If both parts are rigorously demonstrated then the equivalence of the (infinite) time averages and the corressponding phase space averages follows just as in the Landau's reasoning because this then means that the distribution function $\rho(q, p)$ defined as in (225) is uniform.

The proof of the first part is simple. Consider a set of points in the phase space which at t_1 fill the domain $\Delta\Gamma_{(q,p)}(t_1)$ as in Figure 22. At a later instant t_2 the same points fill the domain $\Delta\Gamma_{(q,p)}(t_2)$ which is uniquely determined by the positions at t_2 of those phase points which at t_1 formed the boundary of $\Delta\Gamma_{(q,p)}(t_1)$. By the Liouville theorem the volumes of the two domains are equal (though their shapes may be different). Let us follow two boundary points, A and B, which lie on the same trajectory as in Figure 22. The time t_{AC} it takes the point A to reach C must be therefore the same as the time t_{BD} it takes the point B to reach D because these two points and their images define the boundaries of $\Delta\Gamma_{(q,p)}(t_1)$ and $\Delta\Gamma_{(q,p)}(t_2)$. Moreover

$$t_{AC} = t_{AB} + t_{BC}\,, \qquad t_{BD} = t_{BC} + t_{CD}\,,$$

from which it follows, because $t_{AC} = t_{BD}$ by assumption, that $t_{AB} = t_{CD}$: the phase point does indeed spend the same times in $\Delta\Gamma_{(q,p)}(t_1)$ and $\Delta\Gamma_{(q,p)}(t_2)$ of equal volumes. Furthermore, if τ is sent to infinity, any time the phase point passes through the domain $\Delta\Gamma_{(q,p)}(t_1)$ it must also pass through $\Delta\Gamma_{(q,p)}(t_2)$.

The difficulty lies in showing that this result holds wherever the volumes are located. The original hypothesis of Boltzmann (called by him the ergodic hypothesis) was that as $\tau \to \infty$ the trajectory passes through every point of the constant energy hypersurface. As this was untenable on mathematical grounds, it has been replaced by the so-called quasi-ergodic hypothesis (now frequently called just ergodic) that as $\tau \to \infty$ the trajectory passes through any arbitrarily small neighbourhood of every point of the constant energy hypersurface. Although this may seem physically obvious if the system is large, it is not easy to prove mathematically. The important step towards the complete proof was done by Birkhoff. The crucial problem is in showing that the constant energy hypersurface does not split into parts mutually inaccessible from one another (the so called metric indecomposability of this hypersurface). Some results have been reached by mathematicians (Sinai) in this direction but concerning idealized systems like the model of N hard spheres (elastically scattering on one another). No general proof exists.

At present the common view is that, while being very interestig as a problem in pure mathematics and also as an important part of the investigations of chaotic phenomena exhibited by complex classical systems, the ergodic problem is not very relevant to the physical foundations of classical statistical mechanics and we mentioned it here only in order to make students (with more mathematical inclinations) aware of its existence as a potentially interesting field of research and to at least partly remove an atmosphere of misticism surrounding it.

LECTURE XI (STAT)

Quantum mechanics
As it should be known, the primitive notion of quantum mechanics is the one of the *quantum state* of the considered system.[156] In the mathematical formalism of quantum mechanics quantum states are represented in a Hilbert space $\mathcal{H}$ - a complete (in the sense of convergence in it of all Cauchy sequences) vector space over the field of complex numbers endowed with a scalar product $(\cdot|\cdot)$ and therefore also a norm $\| \cdot \|$ - by classes (called rays) of equivalence of vectors: two vectors Ψ and $\Psi' = e^{i\delta}\Psi$ belong to the same ray (and represent the same physical state of the system). The proper Hilbert space $\mathcal{H}$ consists of normalizable vectors, i.e. such vectors that $\| \Psi \|^2 = (\Psi|\Psi) < \infty$, but in many cases it is convenient to enlarge it (essentially replacing it with its dual $\mathcal{H}^*$ - the space of linear forms over $\mathcal{H}$ - and identifying elements of the proper Hilbert space with the one-forms corresponding to them via the Frechet-Riesz isomorphism) including also the nonnormalizable vectors (also called generalized vectors). Representants of rays are called state-vectors. Observables - quantities which can be measured on the system - are represented by linear self-adjoint operators $\hat{O}$ (also called - not entirely correctly - Hermitian operators), that is such that $\hat{O}^\dagger = \hat{O}$, where the operator $\hat{O}^\dagger$ adjoint with respect to $\hat{O}$ is defined by the equality

$$(\Phi|\hat{O}^\dagger\Psi) = (\hat{O}\Phi|\Psi)\,, \tag{244}$$

which must hold for all vectors Ψ and Φ belonging[157] to $\mathcal{H}$. The spectrum - the set of its eigenvalues o_k i.e. the set of such numbers that there exists vectors satisfying the equation $\hat{O}\Psi_k = o_k\Psi_k$ - of a self adjoint operator $\hat{O}$ is real which enables one to identify it with possible outputs of individual measurements on the system of the physical quantity represented by this operator. The most important operator for the quantum mechanics of a given physical system is the (self-adjoint) Hamiltonian $\hat{\mathscr{H}}$ representing the (total) energy of this system. It determines the system's state time evolution via the Schrödinger equation

$$i\hbar\,\frac{d}{dt}\,\Psi(t) = \hat{\mathscr{H}}\Psi(t)\,. \tag{245}$$

---

[156] Let us remark that to some this notion remains still too abstract to be immediately accepted as a basis of the whole quantum mechanics. In connection with this it is amusing to read the first chapter of the Dirac's Principles of Quantum Mechanics and then the preface written to the russian translation of this renowned monograph by the academician V.I. Fock: the preface clearly shows that its author, who himself made a significant contribution to the quantum theory - we will use in this Course the Fock space, when we come to discuss systems exchanging matter with the surrounding - evidently could not liberate himself from the conventional notion of the "wave function" understood narrowly in the spirit of the Schrödinger wave mechanics.

[157] We are slightly simplifying things here in order not to enter into somewhat subtle but largely irrelevant for us issues related to domains of operators and distinctions between Hermitian (symmetric) and truly self-adjoint operators.

If the Hamiltonian operator is not explicitly time-dependent, the solution of the Schrödinger equation is given by

$$\Psi(t) = \exp\left(-\frac{i}{\hbar}\,\hat{\mathscr{H}}\,(t-t_0)\right)\Psi(t_0)\,, \tag{246}$$

with the state-vector $\Psi(t_0)$ playing the role of the initial conditions. The spectrum of the Hamiltonian which we will assume (as is appropriate for finite systems usually considered in statistical mechanics) to be purely discrete (but which, in general, if quantum mechanics is formulated in the infinite space, can consist also of a continuos part - the corresponding eigenvectors are nonnormalizable generalized vectors belonging to $\mathcal{H}^*$) constitutes always the most important characteristics of the considered physical system.

Wherever more subtle mathematical issues do not intervene, convenient is the Dirac notation in which state-vectors are written as "kets" $|\Psi\rangle$ and elements of the dual space as $\langle\Phi|$. The scalar product $(\Phi|\Psi)$ takes in this notation the form (which in fact implicitly employs the Frechet-Riesz isomorphism) $\langle\Phi|\Psi\rangle$ and the action of an operator $\hat{O}$ on a state vector Ψ is written as $\hat{O}|\Psi\rangle$. The quantities $\langle\Phi|\hat{O}|\Psi\rangle$ are called matrix elements[158] of the operator $\hat{O}$ between the states Ψ and Φ.

As is usually the case with vector spaces, also in the Hilbert space it is possible to choose a basis, a set of vectors Ψ_l, or just $|l\rangle$, labeled by some label l; the basis of the proper Hilbert space $\mathcal{H}$ can always be chosen so that the label l, which can also stand for a multi-label of the sort $l_1l_2\dots l_n$, is discrete. Only calculational convenience dictates employing continuos labels (the corresponding vectors are then generalized nonnormalizable vectors, i.e. elements of $\mathcal{H}^*$). If a countable basis can be chosen (so that l runs over a finite or a countably infinite set of values), the Hilberst space $\mathcal{H}$ is separable. If it is uncountable - e.g. when $l = l_1l_2\dots l_n$ with $n = \infty$ and each l_i runs over a countably infinte set - then $\mathcal{H}$ is nonseparable and infinitely many mutually orthogonal separable subspaces can be chosen in it; in statistical physics nonseparable Hilbert spaces enter the game with the Grand Canonical Ensemble and lie in fact at the basis of the theoretical description of several interesting phenomena (Bose-Einstein condenstation, phase transitions, which can be studied only after the thermodynamic limit (242) is properly taken), but except for a single digression in Lecture XIII we will not enter into these fine mathematical details in this Course. Any state-vector can be written as a linear combination

$$\Phi = \sum_l \Psi_l\, c_l\,, \qquad \text{or} \qquad |\Phi\rangle = \sum_l |l\rangle\, c_l\,, \tag{247}$$

of vectors Ψ_l (or $|l\rangle$) forming the chosen basis. Usually one works with bases formed by mutually orthonormal vectors: $(\Psi_{l'}|\Psi_l) = \langle l'|l\rangle = \delta_{l'l}$, which are eigenvectors of an observable $\hat{O}$ (in most cases of the Hamiltonian $\hat{\mathscr{H}}$ of the system under consideration) which can be constructed using the well known Gramm - Schmidt orthonormalization prescription.

---

[158] The Dirac notation is adapted to self adjoint operators; if $\hat{O}$ is not self adjoint one has to adopt the convention that the symbol $\langle\Phi|\hat{O}|\Psi\rangle$ means $(\Phi|\hat{O}\Psi)$. The scalar product $(\hat{O}\Phi|\Psi)$ can in this notation be written only as $(\langle\Psi|\hat{O}|\Phi\rangle)^*$.

Matrix elements of an operator $\hat{O}$ between the state-vectors $|l\rangle$ of the orthonormal basis formed by its own eigenvectors are "diagonal" that is, take the form

$$\langle l'|\hat{O}|l\rangle = o_l\,\delta_{l'l}\,, \tag{248}$$

and the operator itself can be written in the form

$$\hat{O} = \sum_l |l\rangle\, o_l\, \langle l| = \sum_{o_l\in(\text{spectrum})} o_l\,\hat{P}_{o_l}\,, \tag{249}$$

where $\hat{P}_{o_l} = \hat{P}^2_{o_l}$ is the projector onto the subspace (which may well be multidimensional) of $\mathcal{H}$ corresponding to the eigenvalue o_l of the operator $\hat{O}$. Also important is the fact that if two operators $\hat{O}_1$ and $\hat{O}_2$ commute, that is $[\hat{O}_1,\,\hat{O}_2] \equiv \hat{O}_1\hat{O}_2 - \hat{O}_2\hat{O}_1 = 0$, it is possible to find the basis in which both these operators are simultaneously diagonal.

Formulation of quantum statistical mechanics
In the previous Lecture we have formulated the approach to classical statistical mechanics based on the use of ensembles in general terms so as to make it applicable in principle also to systems not in equilibrium. Only later we have narrowed down the discussion to systems in equilibrium (either macroscopically isolated or in equilibrium with their surrounding). Our discussion of statistical methods based on quantum mechanics will be from the beginning restricted to systems in equilibrium because I'm not very familiar (although I wish I would be!) with the treatment of nonequilibrium quantum systems. (But most probably it the fomulation given below can be straightforwardly extended to systems not in equilibrium).

There are two circumstances which make the assumption that systems macroscopically looking as perfectly isolated cannot be treated as such at the microscopic level even more true in the quantum case than in classical physics. The first one is the extreme density of the spectra of Hamiltonians of real quantum systems which are macroscopically large (i.e. consist of $N \sim N_A$ particles or molecules or other elements). Statistical systems are always assumed to be of finite (although macroscopically large) spatial dimensions and therefore the problems associated with nonnormalizable (generalized) eigenvectors of Hamiltonians and their related continuous (parts of) energy spectra can in principle always be avoided.[159] Yet the spectra of Hamiltonians of macroscopic systems are so

---

[159] This in fact can be said of any physical system which can be of interest (barring somewhat abstract questions about the quantum state of the entire Universe). However, there is a strong physical conviction that local physical results cannot depend on whether the theory is formulated in the continuum (infinite space) or in a finite (but sufficiently large) volume. Even problems like the one of the Hydrogen atom energy spectrum should not depend on this, at least from the practical point of view, eventhough in the finite volume the entire spectrum is discrete and the density of states near $E = 0^-$ and above zero is significantly modified. (Also in many problems in my own favourite branch of theoretical physics - in quantum field theory - it is good to remember that in fact the theory should be formulated in a finite space.) From the computational point of view formulation of the theory in the infinite space has many advantages. The problem however is that quantum mechanics in a finite volume and in the continuum are mathematically very different! To give the simplest example: the scattering theory which can be used

dense that they are nearly continuous and in many situations can formally be treated as such. Roughly, the energy levels of such systems are separated by gaps of order

$$E_{n+1} - E_n \sim \exp(-N)\,.$$

This character of the spectrum is not seen if the N elements of which the system is composed are mutually noininteracting (examples used in introductory courses of statistical physics are mostly of this kind) - in such cases it is the degeneracy of individual energy levels which grows exponentially with N (this will be illustrated in classes by the system of N non-interacting particles confined in a macroscopic finite volume V) - but mutual interactions, even if they can be neglected in computing gross macroscopic features of the system, must always be present (otherwise the system could not reach equilibrium) and always cause fine spliting of degenerate energy levels so that the above estimate of typical interlevel gap becomes true. The estimate remains true also in the case of strongly interacting systems, when the energy levels of the Hamiltonians cannot be seen as perturbations of the free Hamiltonian spectrum which can be interpreted in terms of individual energy levels of separate particles (elements) composing the system. In view of this extreme narrowness of gaps between energy levels of a macroscopic system, any external perturbation, however weak, is associated with an energy transfer to or from the system which is much much larger than its energy gaps and therefore the system, even if it macroscopically perceived as isolated, is continuously making innumerable transitions between its different energy eigenstates.

The second circumstance has its origin in the energy-time uncertainty principle which tells that preparing a quantum system in a state in which its energy uncertainty is ΔE takes at least a time Δt related to ΔE by

$$\Delta E\,\Delta t \gtrsim \hbar\,.$$

In view of the density of energy levels of a macroscopic body (discussed above) preparing the system in a states of definite energy (i.e. ΔE smaller than typical energy splittings) woud require a time longer than the lifetime od the Universe.[160] For both these reasons one should treat all macroscopic quantum systems, even the macroscopically isolated ones, as interacting with their surrounding and certainly not in a stationary state of definite energy. Therefore we shall now develop the formalism of the density operator and of mixed states which allows to treat such macroscopic systems.

Mixed states and the density operator

Suppose the world is divided into the considered system and its surrounding. Even if states of elements which constitute the system and those which constitute the surrounding can be

---

if the quantum mechanics of a single particle moving in a potential in one dimension is formulated in the continuum does not exist if the same problem is formulated on a finite segment of $\mathbb{R}^1$, no matter how large, and all quantities like transmition and reflection coefficients must be replaced by other characteristics simply because the asymptotic (scattering) states cannot be in this case defined.

[160] Although in equilibrium statistical mechanics time does not play any role, its formalism is supposed to apply to real systems the preparation time of which certainly can not be infinite.

separately identified and therefore the Hilbert space $\mathcal{H}$ of the entire universe (consisting of the system and the surrounding) can be represented as the tensor product $\mathcal{H} = \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{env}}$ which in particular means that as its basis one can take vectors

$$|l, L\rangle \equiv |l\rangle \otimes |L\rangle \,, \tag{250}$$

where $|l\rangle \in \mathcal{H}_{\text{sys}}$ and $|L\rangle \in \mathcal{H}_{\text{env}}$, quantum states of both, of the system and of the surrounding, i.e. of the entire content of the Universe, are usually highly entangled, that is, state-vectors representing them in $\mathcal{H}$ must be written as general superpositions

$$|\Psi\rangle = \sum_l \sum_L |l\rangle \otimes |L\rangle \, c_{L,l} \,, \tag{251}$$

and in general cannot be factorized, i.e. are not of the form $|\psi_{\text{sys}}\rangle \otimes |\phi_{\text{env}}\rangle$ (the coefficients $c_{L,l}$ are not in general of the form $c_{L,l} = c_L^{\text{env}} c_l^{\text{sys}}$).

Suppose now that in the Hamiltonian of the Universe $\hat{\mathscr{H}} = \hat{\mathscr{H}}_{\text{sys}} + \hat{\mathscr{H}}_{\text{env}} + \hat{\mathscr{H}}_{\text{int}}$ in which $\hat{\mathscr{H}}_{\text{sys}}$ stands for $\hat{\mathscr{H}}_{\text{sys}} \otimes \hat{1}_{\text{env}}$ and $\hat{\mathscr{H}}_{\text{env}}$ for $\hat{1}_{\text{sys}} \otimes \hat{\mathscr{H}}_{\text{env}}$ (which means that each of these operators acts essentially in only one space of the tensor product[161]) the interaction term is small and we are interested in an observable pertaining to the system alone and not to its surrounding. Such an observable is represented, therefore, by an operator of the form $\hat{O} \equiv \hat{O}_{\text{sys}} \otimes \hat{1}_{\text{env}}$. We would like to express its quantum mechanical expectation value in a state of the form (251) entirely by objects pertaining to $\mathcal{H}_{\text{sys}}$. This can be done with the help of the density operator of the (real) system which is introduced in the following way. We start by writing

$$\begin{aligned}\langle\Psi|\hat{O}|\Psi\rangle &= \left(\sum_{l',L'} c^*_{L',l'} \langle L'| \otimes \langle l'|\right) \hat{O}_{\text{sys}} \otimes \hat{1}_{\text{env}} \left(\sum_{l,L} |l\rangle \otimes |L\rangle \, c_{L,l}\right) \\ &= \sum_{l',l} \langle l'|\hat{O}_{\text{sys}}|l\rangle \sum_L c_{L,l} \, c^*_{L,l'} \,. \end{aligned} \tag{252}$$

We have used the way $\hat{O}$ acts on the state-vectors of the product basis of the complete Hilbert space $\mathcal{H}$:

$$\hat{O}(|l\rangle \otimes |L\rangle) = \hat{O}_{\text{sys}} \otimes \hat{1}_{\text{env}}(|l\rangle \otimes |L\rangle) = (\hat{O}_{\text{sys}}|l\rangle) \otimes (\hat{1}_{\text{env}}|L\rangle) = (\hat{O}_{\text{sys}}|l\rangle) \otimes |L\rangle \,,$$

and the orthogonality[162] $\langle L'|L\rangle = \delta_{L'L}$ of the basis of $\mathcal{H}_{\text{env}}$. The **density operator** $\hat{\rho}$ acting in $\mathcal{H}_{\text{sys}}$ (which depends, of course, on the state of the system's surrounding and

---

[161] An operator of the form $\hat{A}_1 \otimes \hat{A}_2$ acts on a vector of the form $|l_1\rangle \otimes |l_2\rangle$ belonging to the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ of two Hilbert spaces according to the rule

$$\hat{A}_1 \otimes \hat{A}_2 \left(|l_1\rangle \otimes |l_2\rangle\right) = \left(\hat{A}_1|l_1\rangle\right) \otimes \left(\hat{A}_2|l_2\rangle\right) \,,$$

and its action is extended to the entire $\mathcal{H}_1 \otimes \mathcal{H}_2$ by linearity.

[162] The scalar product of the vectors $|\psi_{\text{sys}}\rangle \otimes |\phi_{\text{env}}\rangle$ and $|\psi'_{\text{sys}}\rangle \otimes |\phi'_{\text{env}}\rangle$ belonging to $\mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{env}}$ is naturally defined as

$$\left(\langle\phi'_{\text{env}}| \otimes \langle\psi'_{\text{sys}}|\right) \left(|\psi_{\text{sys}}\rangle \otimes |\phi_{\text{env}}\rangle\right) = \langle\phi'_{\text{env}}|\phi_{\text{env}}\rangle \cdot \langle\psi'_{\text{sys}}|\psi_{\text{sys}}\rangle \,,$$

and extended using linearity to all vectors of $\mathcal{H} = \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{env}}$.

evolving in time with it) is now defined by giving its matrix elements between the basis vectors $|l\rangle$ of $\mathcal{H}_{\rm sys}$:

$$(\Psi_l|\hat{\rho}\,\Psi_{l'}) \equiv \langle l|\hat{\rho}|l'\rangle = \sum_L c_{L,l}\,c^*_{L,l'}\,. \tag{253}$$

Using this operator the expectation value (252) can be written in the form

$$\langle\Psi|\hat{O}|\Psi\rangle = \sum_{l'l}\langle l|\hat{\rho}|l'\rangle\langle l'|\hat{O}_{\rm sys}|l\rangle = {\rm Tr}(\hat{\rho}\,\hat{O}_{\rm sys})\,, \tag{254}$$

in which the trace on the right hand side is restricted to $\mathcal{H}_{\rm sys}$.

The operator $\hat{\rho}$ acting in $\mathcal{H}_{\rm sys}$ defined by the equality (253) is Hermitian (self-adjoint). To see this, it suffices to apply the definition (244) to the matrix elements of $\hat{\rho}^\dagger$ between the basis vectors $|l\rangle$, or Ψ_l:

$$(\Psi_{l'}|\hat{\rho}^\dagger\,\Psi_l) = (\hat{\rho}\,\Psi_{l'}|\Psi_l) = (\Psi_l|\hat{\rho}\,\Psi_{l'})^* = \left(\sum_L c_{L,l}c^*_{L,l'}\right)^* = \sum_L c^*_{L,l}c_{L,l'}\,,$$

and to compare this with the matrix elements of $\hat{\rho}$ given by (253):

$$(\Psi_{l'}|\hat{\rho}\,\Psi_l) = \langle l'|\hat{\rho}|l\rangle = \sum_L c^*_{L,l}c_{L,l'}\,.$$

The two right hand sides are identical, which shows that $\hat{\rho}^\dagger = \hat{\rho}$.

As every self-adjoint operator, $\hat{\rho}$ can be diagonalized, that is there exists in $\mathcal{H}_{\rm sys}$ a complete set of orthonormal basis vectors $\Phi_k \equiv |k\rangle$ which are eigenvectors of $\hat{\rho}$ and the corresponding eigenvalues w_k are real. In this basis

$$(\Phi_k|\hat{\rho}\,\Phi_{k'}) = \langle k|\hat{\rho}|k'\rangle = w_k\,\delta_{k'k} \tag{255}$$

If one now takes the expectation value (252) of the unit operator $\hat{1} = \hat{1}_{\rm sys}\otimes\hat{1}_{\rm env}$ so that obviously $\langle\Psi|\hat{1}|\Psi\rangle = 1$ and writes the right hand side of the formula (254) using the vectors $|k\rangle$ as the basis of $\mathcal{H}_{\rm sys}$ (instead of $|l\rangle$'s), one will get the equality

$$1 = \langle\Psi|\hat{1}|\Psi\rangle = {\rm Tr}(\hat{\rho}) = \sum_k\langle k|\hat{\rho}|k\rangle = \sum_k w_k\,. \tag{256}$$

Thus, the sum of the eigenvalues of the statistical operator equals one. Another information about the eigenvalues w_k can be obtained by realizing that $\hat{\rho}$ can be written as[163]

$$\hat{\rho} = \sum_k |k\rangle w_k\langle k|\,, \tag{257}$$

---

[163] Indeed,

$$\langle k''|\hat{\rho}|k'\rangle = \langle k''|\left(\sum_k |k\rangle\,w_k\,\langle k|\right)|k'\rangle = \sum_k\langle k''|k\rangle\,w_k\,\langle k|k'\rangle = w_{k'}\,\delta_{k'k''}\,.$$

and using this representation to write in two ways $\mathrm{Tr}(\hat{\rho}\,\hat{O}_{\mathrm{sys}})$ with the particular operator $\hat{O}_{\mathrm{sys}} = |k\rangle\langle k|$, i.e. with a self-adjoint projector onto the $\hat{\rho}$ eigensubspace corresponding to the eigenvalue w_k. On one hand then

$$\mathrm{Tr}(\hat{\rho}\,\hat{O}_{\mathrm{sys}}) = \sum_{k''}\langle k''| \left(\sum_{k'} |k'\rangle\, w_{k'}\, \langle k'|\right) (|k\rangle\langle k|)\, |k''\rangle = w_k\,,$$

and on the other, using the definition (253),

$$\mathrm{Tr}(\hat{\rho}\,\hat{O}_{\mathrm{sys}}) = \sum_{l',l}\langle l|\hat{\rho}|l'\rangle\langle l'|\hat{O}_{\mathrm{sys}}|l\rangle = \sum_{L}\sum_{l',l} c_{L,l}\, c^{*}_{L,l'}\, \langle l'|k\rangle\langle k|l\rangle = \sum_{L}\left|\sum_{l} c_{L,l}\langle k|l\rangle\right|^2 \geq 0\,.$$

Thus, $0 \leq w_k \leq 1$. Finally, written in the basis of the density operator eigenvectors $|k\rangle$ the formula (254) takes the form

$$\langle\Psi|\hat{O}|\Psi\rangle = \mathrm{Tr}(\hat{\rho}\,\hat{O}_{\mathrm{sys}}) = \sum_{k} w_k\, \langle k|\hat{O}_{\mathrm{sys}}|k\rangle\,. \tag{258}$$

It follows that if the system interacting with its surrounding is considered separately, it cannot be viewed as being in a well-defined quantum state, called **pure** state, which in the Hilbert space can be represented by a state-vector which can always be written as a superposition (however complicated) of state-vectors of some basis. Instead, its state must be represented by a density operator $\hat{\rho}$; the system is then said to be in a **mixed state**. The difference between the two situations is best seen by writing a pure state $|\psi\rangle$ (of the system) as a superposition of the basis state-vectors $|k\rangle$ of $\mathcal{H}_{\mathrm{sys}}$ which are eigenvectors of $\hat{\rho}$. The expectation value of an operator $\hat{O}_{\mathrm{sys}}$ in the pure state $|\psi\rangle = \sum_k |k\rangle c_k$ reads

$$\langle\psi|\hat{O}_{\mathrm{sys}}|\psi\rangle = \sum_{k'k} c^{*}_{k'} c_k\, \langle k'|\hat{O}_{\mathrm{sys}}|k\rangle = \sum_{k} |c_k|^2\langle k|\hat{O}_{\mathrm{sys}}|k\rangle + \sum_{k'\neq k} c^{*}_{k'} c_k\, \langle k'|\hat{O}_{\mathrm{sys}}|k\rangle\,. \tag{259}$$

The first term resembles the formula (258) if the factors $|c_k|^2$ are identified[164] with w_k. But the expectation value of $\hat{O}_{\mathrm{sys}}$ in the pure state $|\psi\rangle$ involves also the second term which has no counterpart in (258). One may say that while a pure state $|\psi\rangle$ is always a **coherent** superposition of the basis states $|k\rangle$ with definite relative phases φ_k of the superposition coefficients c_k, a mixed state can be viewed as their **incoherent** superposition, with random relative phases of the coefficients of the superposition, because (259) goes over into (258) upon averaging (with the flat probability distribution in the $[0, 2\pi]$ range) over the phases φ_k of the coefficients $c_k = |c_k| \exp(i\varphi_k)$ (the products $c^{*}_{k'} c_k$ with $k \neq k'$ are removed by such averaging as a result of the statistical independence of different phases). Mixed states are generalization of pure states because every pure state can be represented in the formalism of mixed states by the density operator $\hat{\rho}$ satisfying the equality

$$\hat{\rho}^2 = \hat{\rho}\,, \tag{260}$$

---

[164] Similarly to the w_k's, $0 \leq |c_k|^2 \leq 1$.

which means that in the basis formed by its eigenvectors the operator $\hat{\rho}$ is such that $w_k^2 = w_k$ and since the factors w_k are nonnegative and must sum up to unity, this implies that only one $w_k = 1$ and all others are zero. In other words, in the formalism of mixed states, a pure state normally represented by the normalized to unity state-vector $|\psi\rangle$ is represented by the density operator $\hat{\rho} = |\psi\rangle\langle\psi|$ which is simply the projection operator onto the subspace spanned in the Hilbert space by this vector (it is evident that it is in fact the ray, and not the state-vector itself, which determines $\hat{\rho}$).

Ensembles in quantum statistical mechanics
Thus any real system, if considered separately, must be at a given instant t_0 treated as being in a mixed state represented by some density operator $\hat{\rho}$ which is diagonal in some basis $|k\rangle$ and takes the form (257) at some initial moment t_0. Of course, as the state of the entire universe evolves with time, the density operator $\hat{\rho}$ representing the considered real system in $\mathcal{H}_{\rm sys}$ changes too and, as in the classical case, these changes cannot be determined by the Hamiltonian of the system alone; this would require following the time evolution of the state of the entire universe.

As in the classical case, in view of this situation, one resorts to the method of statistical ensembles. The statistical ensemble corresponding to a real system microscopically perturbed by the rest of the universe is a collection of $\mathcal{N}$ absolutely **isolated systems** each of which is at the initial instant t_0 in one of the pure states $|k\rangle$. To be representative for the considered real system (which may be macroscopically isolated or in equilibrium with its surrounding), the relative numbers $\mathcal{N}_k/\mathcal{N}$ of systems in different pure states $|k\rangle$ in the ensemble at the initial moment t_0 (which, because the systems in the ensemble are isolated, can be taken to be $t_0 = 0$) must be such that (the limits $\mathcal{N}_k \to \infty$, $\mathcal{N} \to \infty$ are understood)

$$\mathcal{N}_k/\mathcal{N} = w_k\,, \qquad \text{where} \qquad \mathcal{N} = \sum_k \mathcal{N}_k\,.$$

The (statistical) mean values over the ensemble of expectation values of operators representing observables are then formally identical to the formula (258):

$$\overline{O} = \frac{1}{\mathcal{N}}\sum_k \mathcal{N}_k\langle k|\hat{O}|k\rangle = \sum_k w_k\langle k|\hat{O}|k\rangle\,, \tag{261}$$

and can be written as

$$\overline{O} = \mathrm{Tr}(\hat{\rho}\,\hat{O}_{\rm sys})\,, \tag{262}$$

with the **statistical operator**

$$\hat{\rho} = \sum_k |k\rangle\,(\mathcal{N}_k/\mathcal{N})\,\langle k|\,, \tag{263}$$

which at t_0 is formally identical with the density operator $\hat{\rho} = \sum_k |k\rangle w_k\langle k|$ of the real system at that instant.[165] The operator $\hat{\rho}(t)$ replaces the distribution function $\rho(q,p,t)$

---

[165] Because of this formal identity one can take the position that the ensemble is just a single system in an appropriate mixed state.

of classical statistical mechanics. The central problem, as in the classical case, is to theoretically determine the form of the statistical operator which would represent the mixed state of the real system of interest in different situations (when the real system is macroscopically isolated, or is in thermal equlibrium with its surrounding, or in a nonequilibrium state).

Because the systems forming the ensemble are isolated (autonomous), their time evolution is completely determined by the Hamiltonian $\hat{\mathscr{H}}$ (from now on we drop the subscript "sys") which acts only in the system's Hilbert space. Each of the basis state-vectors $|k\rangle$ changes with time according to (246):

$$|k(t)\rangle = \exp(-(i/\hbar)\hat{\mathscr{H}}t)|k\rangle\,.$$

as a result the time evolution of the statistical operator (263) of the ensemble (but not the density operator of the real system!) is

$$\hat{\rho}(t) = \sum_k w_k\, e^{-\frac{i}{\hbar}\hat{\mathscr{H}}t}|k\rangle\langle k|e^{+\frac{i}{\hbar}\hat{\mathscr{H}}t}\,.$$

Differentiating both sides with respect to t gives the differential equation

$$\frac{d\hat{\rho}(t)}{dt} = -\frac{i}{\hbar}\,[\hat{\mathscr{H}},\,\hat{\rho}(t)]\,, \tag{264}$$

satisfied by the ensemble's statistical operator $\hat{\rho}(t)$. This equation plays in quantum statistical mechanics the same role as does the Liouville equation

$$\frac{\partial\hat{\rho}(q,p,t)}{\partial t} = \{\mathscr{H}(q,p),\,\rho(q,p,t)\}_{\rm PB}\,, \tag{265}$$

in classical statistical physics.[166]

If the considered real system is in equilibrium (either as a macroscopically isolated system or as a system in equilibrium with its surrounding), which means that macroscopic

---

[166] The equation (264) looks very similar to the Heisenberg equation

$$\frac{d\hat{O}_H(t)}{dt} = \frac{i}{\hbar}\,[\hat{\mathscr{H}},\,\hat{O}_H(t)]\,,$$

satisfied by time-dependent Heisenberg picture operators $\hat{O}_H(t)$ which are couterparts of time-independent Schrödinger picture operators (for more details, see Chapter 1 of my notes to quantum field theory), but the signs of the right hand sides are different; in fact the Heisenberg picture counterpart $\hat{\rho}_H$ of the statistical operator $\hat{\rho}(t)$ is time independent. It is also worth noting that the transition from the classical Liouville equation (265) to the equation (264) can be done with the usual "quantization rule" according to which in classical equations all c-number functions are replaced by their operator counterparts and the Poisson brackets are replaced by commutators according to the prescription

$$\{\cdot\,,\,\cdot\}_{\rm PB} \to -\frac{i}{\hbar}\,[\cdot\,,\,\cdot]\,.$$

observables measured on it do not depend on time, the statistical operator of the representative ensemble should be time independent, just as was the distribution functions $\rho(q, p)$ corresponding to a classical system in equilibrium (only then the ensemble mean values (262) will be stationary), From the equation (264) it follows that $d\hat{\rho}/dt = 0$ implies

$$[\hat{\mathscr{H}}, \hat{\rho}] = 0 \,. \tag{266}$$

According to the usual rules of quantum mechanics this means that there exist in $\mathcal{H}$ a basis $|n\rangle$ in which both operators, $\hat{\mathscr{H}}$ and $\hat{\rho}$, are simultaneously diagonal:

$$\hat{\mathscr{H}} = \sum_n E_n |n\rangle\langle n| \,, \qquad \hat{\rho} = \sum_n w_n |n\rangle\langle n| \,. \tag{267}$$

Another information on the form of the statistical operator corresponding to a macroscopic system in equilibrium is provided by the statistical independence of its macroscopic subsystems. If the system is composed of two[167] mutually nointeracting, or interacting negligibly weakly, parts, its Hilbert space $\mathcal{H}$ can be taken to be the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and the system's Hamiltonian is $\hat{\mathscr{H}} = \hat{\mathscr{H}}_1 \otimes \hat{1}_2 + \hat{1}_1 \otimes \hat{\mathscr{H}}_2$ (the term $\hat{\mathscr{H}}_{\rm int}$ being either absent or negligible). The basis of $\mathcal{H}$ built out of the Hamiltonian eigenvectors $|n\rangle$ can be then chosen to be of the form

$$|n\rangle \equiv |n_1, n_2\rangle \equiv |n_1\rangle \otimes |n_2\rangle \,,$$

The statistical operator of the corresponding ensemble should then have the form $\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$ which ensures that the mean value $\overline{O}$ of an observable represented by the operator of the form $\hat{O}_1 \otimes \hat{O}_2$, that is which is a product of observables pertaining to different systems, equals $\overline{O_1}\,\overline{O_2}$:

$$\begin{aligned}
\overline{O} = \overline{O_1 O_2} &= {\rm Tr}_{\mathcal{H}}\Big(\hat{\rho}_1 \otimes \hat{\rho}_2\, \hat{O}_1 \otimes \hat{O}_2\Big) \\
&= {\rm Tr}_{\mathcal{H}}\Big(\hat{\rho}_1\hat{O}_1 \otimes \hat{\rho}_2\hat{O}_2\Big) = \sum_{n_1 n_2} \langle n_2, n_1|\hat{\rho}_1\hat{O}_1 \otimes \hat{\rho}_2\hat{O}_2|n_1, n_2\rangle \\
&= \sum_{n_1} \langle n_1|\hat{\rho}_1\,\hat{O}_1|n_1\rangle \sum_{n_2} \langle n_2|\hat{\rho}_2\,\hat{O}_2|n_2\rangle = {\rm Tr}_{\mathcal{H}_1}\Big(\hat{\rho}_1\,\hat{O}_1\Big)\,{\rm Tr}_{\mathcal{H}_2}\Big(\hat{\rho}_2\,\hat{O}_2\Big) = \overline{O_1}\;\overline{O_2} \,.
\end{aligned}$$

The operators $\hat{\rho}_1$ and $\hat{\rho}_2$ must be both Hermitian (the first one in $\mathcal{H}_1$ and the second one in $\mathcal{H}_2$) and both should be diagonal in the bases formed by the eigenvectors $|n_1\rangle$ and $|n_2\rangle$ of the respective Hamiltonians:

$$\begin{aligned}
w_n^{(12)}\delta_{n'n} = \langle n'|\hat{\rho}|n\rangle &= \langle n'_2, n'_1|\hat{\rho}_1 \otimes \hat{\rho}_2|n_1, n_2\rangle \\
&= \langle n'_1|\hat{\rho}_1|n_1\rangle\,\langle n'_2|\hat{\rho}_2|n_2\rangle = w_{n_1}^{(1)}\,\delta_{n'_1 n_1}\,w_{n_2}^{(2)}\,\delta_{n'_2 n_2} \,.
\end{aligned}$$

In other words, $w_n^{(12)} \equiv w_{n_1,n_2}^{(12)} = w_{n_1}^{(1)} w_{n_2}^{(2)}$, or

$$\ln w_n^{(12)} \equiv \ln w_{n_1,n_2}^{(12)} = \ln w_{n_1}^{(1)} + \ln w_{n_2}^{(2)} \,, \tag{268}$$

---

[167] Extension to more than two independent subsystems is straightforward.

which, since energies of the two (noninteracting or interacting negligibly weakly with one another) parts of the system are additive, $E_{n_1n_2} = E_{n_1} + E_{n_2}$ while, in general, the probabilities can be viewed as functions of energies: $w_{n_i} = w(E_{n_i})$, means that

$$\ln w(E_{n_1}) = \alpha_1 - \beta E_{n_1}\,, \qquad \text{and} \qquad \ln w(E_{n_2}) = \alpha_2 - \beta E_{n_2}\,, \tag{269}$$

as this is the only way of satisfying the requirement (268) if $E_n \equiv E_{n_1n_2,\dots} = E_{n_1}+E_{n_2}+\dots$. All this can be concisely written in the form[168]

$$\ln\hat\rho_{(12)} = (\ln\hat\rho_1)\otimes\hat 1_2 + \hat 1_1\otimes(\ln\hat\rho_2)\,. \tag{270}$$

This imposes a strong constraint on statistical operators of macroscopic systems and will be crucial in deriving the general formula for entropy.

The Quantum Microcanical Ensemble

We now consider a real system which macroscopically can be treated as isolated from its surrounding and in equilibrium, or more precisely, macroscopically not exchanging heat with its surrounding and having fixed values (i.e. directly controlled) of its deformative parameters. Again, appealing to the principle of a priori equal probabilities, we postulate that the statistical operator $\hat\rho$ of the representative ensemble (of strictly, that is, also at the microscopic level, isolated systems) has the form

$$\hat\rho = \text{Const.} \sum_{\substack{n \\ E\le E_n\le E+\Delta E}} |n\rangle\langle n|\,, \tag{271}$$

where $|n\rangle$ are the eigenvectors of the ensemble systems Hamiltonian (that is, of the Hamiltonian of the real system from which the terms corresponding to its interaction with the rest of the universe have been removed). In (271) it has been taken into account that as a result of its microscopic residual interaction with the surrounding as well as a consequence of the mentioned restriction imposed by the general quantum mechanical uncertainty principle, the energy of a real macroscopically isolated system can be specified only up to some tolerance ΔE (of course $\Delta E \ll E$, in the case of a large system). Since $\text{Tr}(\hat\rho) = 1$, the constant in the definition (271) must be given by (the elipses stand for other macroscopic deformative variables like magnetization, etc. which possibly can characterize the system)

$$\text{Const.}^{-1} = \begin{pmatrix} \text{number of quantum states} \\ \text{in the interval } [E, E+\Delta E] \end{pmatrix} \equiv \Gamma(E, V, \dots, N, \Delta E)\,. \tag{272}$$

As the energy levels of the system are discrete, and the volume V is finite, the number of states corresponding to the energy interval $[E, E+\Delta E]$ should be finite.[169] Together the

---

[168] Recall that a function of an operator is defined by giving matrix elements of this operator function in the basis in which the operator is diagonal.

[169] Here it is important that the theory is formulated in the finite volume: for instance, to the range $[-\varepsilon, 0]$ (with arbitrarily small positive ε) there correspond infinitely many discrete energy eigenstates of the Hydrogen atom Hamiltonian (recall that $E_{nlm_l,s_z} = -mc^2\alpha^2_{\text{EM}}/2n^2$ and the degeneracy of the energy levels is $2n^2$), but this is true only if the quantum mechanics of the Hydrogen atom is formulated in the infinite space; if the volume of the space is taken finite, the spectrum gets modified and the number of states in the range $[-\varepsilon, 0]$ is finite (I.I. Oppenheim and D.R. Hafeman, *J.Chem.Phys.* **39** (1963), 101).

formulae (271) and (272) define the **quantum microcanonical ensemble** which should be representative for systems macroscopically isolated and in equilibrium (macroscopically not exchanging heat with its surrounding and having fixed values of its deformative parameters). The form (271) of the statistical operator $\hat{\rho}$ is, of course consistent with the general requirements (266) and (269).

As in the case of the classical microcanonical ensemble, the statistical entropy (called also Boltzmann entropy) of a macroscopically isolated system in equilibrium is defined to be

$$S_{\text{stat}} = k_{\text{B}} \ln \Gamma(E, V, \ldots, N, \Delta E) \,. \tag{273}$$

It should be noticed that in contrast to the classical case, the quantity $\Gamma(E, V, \ldots, N, \Delta E)$ is in the quantum microcanonical ensemble defined in absolute terms, without any arbitrary multiplicative constant which implies that entropy is also defined without any arbitrary additive constants (the dependence on ΔE which cannot be avoided for fundamental reasons, drops out in the thermodynamical limit). As has been explained, the arbitrarines in the classical case is removed by requiring that results of the classical and quantum aproaches match in the appropriate limit.

As in the classical case, in addition to $\Gamma(E, V, \ldots, N, \Delta E)$ one can introduce the quantity $\Sigma(E, V, \ldots, N)$ which here is just the number of quantum states of energies $E_n \leq E$ and, owing to the fact that the spectrum of the Hamiltonian of a macroscopic system, while being formally discrete, is almost continuous (and can formally be treated as such), to define the density $\omega(E, V, \ldots, N)$ of states (around $E_n = E$) by

$$\omega(E, V, \ldots, N) = \frac{\partial \Sigma(E, V, \ldots, N)}{\partial E} \,, \tag{274}$$

so that the relation (240) between $\Gamma(E, V, \ldots, N, \Delta E)$, $\omega(E, V, \ldots, N)$ and ΔE holds true also in the quantum case. Then, again as in the classical case, one can define two other statistical entropies replacing in (273) $\Gamma(E, V, \ldots, N, \Delta E)$ either by $\Sigma(E, V, \ldots, N)$ - this is sometimes called the Gibbs entropy - or by $\omega(E, V, \ldots, N)\Delta E$ (another form of the Boltzmann entropy). In the thermodynamical limit, if the energy spectrum of the system is typical, all the three definitions of S_{stat} lead to the same thermodynamical entropy

$$S_{\text{TMD}} = N \lim_{\infty} \left(\frac{S_{\text{stat}}}{N} \right) , \tag{275}$$

or, in cases in which the number of particles cannot be defined (e.g. because the system is a relativistic quantum field, for instance)

$$S_{\text{TMD}} = V \lim_{\infty} \left(\frac{S_{\text{stat}}}{V} \right) . \tag{276}$$

Thermodynamical entropy obtained in this way in the framework of the Microcanonical Ensemble is given as a function of its natural variables (the dependence on ΔE drops out as a result of taking the thermodynamical limit) and constitutes therefore the fundamental relation (in the Callenian sense) characterizing the system and encodes in it complete thermodynamical information about it.

Before we demonstrate that the statistical entropies defined by (273) have (in the thermodynamical limit) the properties of the entropy defined in the framework of the pure thermodynamics, we will consider the problem of defining the statistical entropy of systems in equilibrium (but not necessarily isolated) in general. Owing to the uniform notation introduced, this can be done jointly for the classical and the quantum case.

The "golden formula" for entropy

Defining the statistical entropy - a measure of the microscopic disorder which is associated with a given equilibrium macrostate - by the formulae (241) and (273) in terms of the phase space volume or the number of quantum states in the classical and quantum cases respectively, was possible because the microcanonical ensemble distribution function $\rho(q,p)$ in the classical case is localized (has a finite support in the mathematical language) and in the quantum case the probabilities w_n of only a finite number of states are nonvanishing. If the real system is in equilibrium with its surrounding (e.g. with a heat bath), the distribution function (in the classical case) or the probabilities w_n (in the quantum case) of the statistical ensemble representing the real system do not have this simple property and the quantity $\Gamma(E,V,N,\Delta E)$ employed in the definitions (241) and (273) is ill defined (unlike the functions $\Sigma(E,V,N)$ and $\omega(E,V,N)$ which are always finite but cannot be directly used to define entropy). Consequently one has to invent another measure of disorder in terms of which to define entropy.

In the general case to introduce the analog of the quantity $\Gamma(E,V,N,\Delta E)$, one can resort to the following reasoning (Landau & Lifschitz). One constructs first the distribution function $\rho_E(E)$ of the random variable E - the system's energy - defined on the space of elementary events. In the classical case it is defined using the prescription (226) which here, because when the considered system is in equilibrium $\rho(q,p)=\rho(\mathscr{H}(q,p))$, gives

$$\begin{aligned}\rho_E(E) &= \int d\Gamma_{(q,p)}\,\rho(\mathscr{H}(q,p))\,\delta(\mathscr{H}(q,p)-E)\\ &= \rho(E)\int d\Gamma_{(q,p)}\,\delta(\mathscr{H}(q,p)-E) = \rho(E)\,\omega(E)\,,\end{aligned}\tag{277}$$

upon using the formula (239). In the quantum case, exploiting the quasi-continuous character of the spectra of macroscopic systems, one can define the analogous distribution $\rho_E(E)$ by the equality

$$\rho_E(E)\,dE = \sum_{\substack{n\\ E\le E_n\le E+dE}} w_n\,.\tag{278}$$

Since $\sum_n w_n = 1$, the distribution $\rho_E(E)$ defined in this way is automatically normalized: $\int dE\,\rho_E(E)=1$. On the other hand, again owing to the quasi-continuity of the spectrum, the number of quantum states in the interval $[E, E + dE]$ is equal to $\omega(E)\, dE$, and the probabilities $w_n = w(E_n)$ can be treated as a continuous function $w(E)$ of the system's energy. This allows in the quantum case to write

$$\rho_E(E) = w(E)\, \omega(E)\,. \tag{279}$$

The two expressions: the classical one (277) and the quantum one (279) become therefore identical if $w(E)$ is identified with $\rho(E)$.

If the considered system is large, the distribution $\rho_E(E)$ is (in most cases) nonzero only in the close vicinity of a value E^* which is practically the same as the mean $\overline{E}$ computed using $\rho_E(E)$. In other words, $\rho_E(E)$ has practically a Dirac delta-like peak almost at $E = \overline{E}$. It is therefore possible to characterize the spread of a given equilibrium macrostate of the system over the phase space in the classical case and over the Hamiltonian spectrum in the quantum case in terms of the "width" ΔE of its energy distribution defined by the relation

$$\rho_E(\overline{E})\, \Delta E = 1\,. \tag{280}$$

This allows to define the *effective* number $\Gamma(\overline{E}, V, N, \Delta E)$ of microstates (that is, the number of those microstates which are really relevant in the ensemble) by the equality (in the classical case $w(E)$ should be replaces by $\rho(E)$)

$$1 = w(\overline{E})\, \omega(\overline{E})\, \Delta E \equiv w(\overline{E})\, \Gamma(\overline{E}, V, N, \Delta E)\,, \tag{281}$$

and then the statistical entropy by

$$S_{\rm stat} = k_{\rm B} \ln \Gamma(\overline{E}, V, N, \Delta E) = k_{\rm B} \ln[\omega(\overline{E})\, \Delta E]\,. \tag{282}$$

It is easy to see that in the case of the mirocanonical statistical operator (271) or the distribution function (236) the definitions (273) and (241) coincide with (282). Indeed (to consider only the quantum case), in this case $w(E) = w(\overline{E}) = 1/\Gamma(E, V, N, \Delta E)$, which implies the equality of $\Gamma(\overline{E}, V, N, \Delta E)$ and $\Gamma(E, V, N, \Delta E)$.

Exploiting now the definition (281) of the effective number of microstates, the formula (282) for entropy can be cast into the form

$$S_{\rm stat} = -k_{\rm B} \ln w(\overline{E})\,, \tag{283}$$

It has been argued, however, that statistical independence of subsystems implies that the logarithm of $w(E)$ must be at most a **linear**[170] function of energy, $\ln w(E) = \alpha - \beta E$ ($\ln \rho(\mathscr{H}) = \alpha - \beta \mathscr{H}$, classically). This means that

$$\ln w(\overline{E}) = \overline{\ln w(E)}\,,$$

which allows to write the formula (283) and, hence, the formula (282), in the form

$$S_{\rm stat} = -k_{\rm B} {\rm Tr}(\hat{\rho} \ln \hat{\rho})\,, \tag{284}$$

---

[170] Here linear in the "school" sense.

in the quantum case and as

$$S_{\rm stat} = -k_{\rm B}\int d\Gamma_{(q,p)}\,\rho(q,p)\ln\rho(q,p)\,, \tag{285}$$

in the classical case. It is clear that in the case of the classical microcanonical ensemble, when $\rho(q,p) = {\rm Const.} = 1/\Gamma(E,\Delta E)$, and

$$\int d\Gamma_{(q,p)}\,\rho(q,p)\ln\rho(q,p) = \frac{1}{\Gamma(E,\Delta E)}\ln\frac{1}{\Gamma(E,\Delta E)}\int_{E\leq\mathscr{H}\leq E+\Delta E} d\Gamma_{(q,p)}\,,$$

the formula (285) reduces to (241). An analogous reasoning readily shows that in the quantum case the formula (284) reduces to (273).

Expressions (284) and (285) are the general, truly golden, formulae for entropy valid independently of the ensemble. Their virtue is that they can be (after some reinterpretation) extended also to nonequilibrium situations (when $\hat{\rho}(t)$ of $\rho(q,p,t)$ are not independent of time), although we will not consider this extension in this Course.

LECTURE XII (STAT)

Classical energy equipartition theorem
One general consequence of the classical statistical approach applied to macroscopic systems in equilibrium, which is independent of whether the system is macroscopically isolated or in equilibrium with its surrounding (the corresponding distribution function $\rho(q,p)$ will be introduced shortly), is the equipartition of its mean energy between its microscopic degrees of freedom. Here we demonstrate this (making also the meaning of this statement more precise) using the classical microcanonical ensemble distribution function (236).

We begin by computing the ensemble average

$$\begin{aligned}\overline{q^i\,\frac{\partial\mathscr{H}}{\partial q^j}} &= \frac{1}{\Gamma(E,\Delta E)}\int_{E\leq\mathscr{H}\leq E+\Delta E} d\Gamma_{(q,p)}\, q^i\,\frac{\partial\mathscr{H}}{\partial q^j}\\ &\approx \frac{\Delta E}{\Gamma(E,\Delta E)}\,\frac{\partial}{\partial E}\int_{\mathscr{H}\leq E} d\Gamma_{(q,p)}\, q^i\,\frac{\partial\mathscr{H}}{\partial q^j}\,.\end{aligned}$$

We have used the standard trick

$$\begin{aligned}\int_{E\leq\mathscr{H}\leq E+\Delta E} d\Gamma_{(q,p)}\, f(q,p) &= \int_{\mathscr{H}\leq E+\Delta E} d\Gamma_{(q,p)}\, f(q,p) - \int_{\mathscr{H}\leq E} d\Gamma_{(q,p)}\, f(q,p)\\ &= F(E+\Delta E) - F(E) \approx \Delta E\,\frac{\partial}{\partial E}\,F(E)\,.\end{aligned}$$

This can be also written as

$$\overline{q^i\,\frac{\partial\mathscr{H}}{\partial q^j}} \approx \frac{\Delta E}{\Gamma(E,\Delta E)}\,\frac{\partial}{\partial E}\int_{\mathscr{H}\leq E} d\Gamma_{(q,p)}\left\{\frac{\partial}{\partial q^j}\left[q^i(\mathscr{H}-E)\right] - \delta^i{}_j(\mathscr{H}-E)\right\}.$$

The first term under the integral (the one with the derivative $\partial/\partial q^j$) will then give zero because the integral over dq^j which is part of the measure $d\Gamma_{(q,p)}$ can be written as the boundary term but at the boundary, which is determined by the condition $\mathscr{H} = E$, vanishes the expression under the derivative.[171] The derivative $\partial/\partial E$ can be then put under the integral (again, differentiating with respect to the dependence on E of the integration boundary would give the integrand evaluated at these boundary where it vanishes) leading to the result

$$\overline{q^i\,\frac{\partial\mathscr{H}}{\partial q^j}} \approx \frac{\Delta E}{\Gamma(E,\Delta E)}\int_{\mathscr{H}\leq E} d\Gamma_{(q,p)}\,\delta^i{}_j \approx \frac{\Delta E}{\omega(E)\Delta E}\,\delta^i{}_j\,\Sigma(E) = \frac{\Sigma(E)}{\omega(E)}\,\delta^i{}_j\,.$$

---

[171] More precisely, the domain of the integrations over the variables q^i should be effectively restricted by introducing a smooth boundary potential $\mathscr{V}(q^1,\dots,q^{3N})$ which is almost zero if all q^i's are in the volume V and tends to infinity if any of these variables is outside V; then the limit

$$\mathscr{V} \to \begin{cases} 0 & \text{if all } q^i\in V\\ \infty & \text{otherwise}\end{cases},$$

should be taken.

(The relation (240) has been used here.) Using the relation $\omega(E) = \partial\Sigma(E)/\partial E$, this result can also be written as

$$\overline{q^i \, \frac{\partial \mathscr{H}}{\partial q^j}} = \delta^i{}_j \left(\frac{\partial \ln \Sigma(E)}{\partial E} \right)^{-1} .$$

Since entropy S within the Microcanonical Ensemble can (from the thermodynamical limit point of view) be defined in terms of the logarithm of $\Sigma(E)$, using the general thermodynamical relation $(\partial S/\partial E) = 1/T$, we obtain

$$\overline{q^i \, \frac{\partial \mathscr{H}}{\partial q^j}} = -\overline{q^i \, \dot{p}_j} = k_{\rm B} T \, \delta^i{}_j \, . \tag{286}$$

Exactly analogous computation[172] gives the result

$$\overline{p_i \, \frac{\partial \mathscr{H}}{\partial p_j}} = \overline{p_i \, \dot{q}^j} = k_{\rm B} T \, \delta_i{}^j \, . \tag{287}$$

These two relations can be used in some special situations to obtain predictions of the classical statistical mechanics without complicated calculations. For instance, if the Hamiltonian has a regular form, e.g.

$$\mathscr{H} = \sum_{i=1}^{K} A_i p_i^2 + \sum_{j=1}^{f} B_j q_j^2 \, ,$$

with some constant A_i and B_j, so that

$$\mathscr{H} = \frac{1}{2} \sum_{i=1}^{K} p_j \, \frac{\partial \mathscr{H}}{\partial p_j} + \frac{1}{2} \sum_{i=1}^{f} q^j \, \frac{\partial \mathscr{H}}{\partial q^j} \, ,$$

then the results (286), (287) imply that the mean[173] energy of the system is related to its temperature by

$$U \equiv \overline{\mathscr{H}} = \frac{1}{2} (K + f) k_{\rm B} T \, . \tag{288}$$

This immediately implies that the heat capacity c_v per molecule of a classical monoatomic perfect gas is $\frac{3}{2} k_{\rm B}$, that is $\frac{3}{2} R$ per mole ($f = 0$, $K = 3N$, so three "degrees of freedom" per molecule) and of a two-atomic gas $\frac{7}{2} k_{\rm B}$ ($K = 6N$, $f = N$, so 7 "degrees of freedom" per molecule). Similarly the heat capacity of the crystal lattice of a solid built out of N

---

[172] Actually even less complicated because one does not need to appeal to an artificial boundary potential.

[173] In the case of the microcanonical ensemble representative for a macroscopically isolated system with fixed energy E (within the tolerance ΔE which in the classical case can be set equal zero) the mean energy $\overline{H} \equiv U$ is the same as E; the result is valid however also if the system is in equilibrium but cannot be ascribed a definite energy.

molecules and behaving as $3N$ independent harmonic oscillators[174] should be $C_V = 3Nk_{\rm B}$ ($K = 3N$, $f = 3N$; this is the so-called Dulong-Petit law) irrespectively of the distribution of frequencies of these oscillators. These results are heuristically formulated as the rule that classically each degree of freedom contributes the amount $\frac{1}{2}k_{\rm B}T$ of energy to the mean system's total energy, but it is clear that the implication of the general results (286), (287) for the total energy depend on the precise form of the interactions.[175] It goes without saying, that the predictions of the equipartition theorem for the heat capacities exemplified above (which are clearly at variance with 3TMDL) are experimentally verified only at sufficiently high temperatures, at which genuine quantum effects (except for the ones related to the indistinguishability of identical particles) are unimportant and systems can be treated as classical.

Canonical (Gibbs) Ensemble energy distribution

The Microcanonical Ensemble introduced in the preceding Lectures representative for macroscopically isolated systems is conceptually simple but very inconvenient in practical computations. Much more convenient in this respect is the Canonical Ensemble which formally is representative for a system remaining in thermal equilibrium with its surrounding (i.e. in thermal contact with it through a diathermal wall) modeled by a large (in the limit infinitely large) heat bath of fixed temperature T. Since in most cases, as far as mean values of system's characteristics are concerned (and not their fluctuations), that is from the point of view of thermodynamics, there should be no difference between an isolated system the temperature of which (defined as $(\partial S/\partial U)_{V,N,\dots}$) is T and the same system exchanging energy (in the form of heat) with a heat bath at temperature T, the Canonical Ensemble is in most cases (at least classically) the preferred way to perform actual computations also in the case of macroscopically isolated systems.

The form of the distribution function $\rho(q,p)$ or of the statistical operator $\hat{\rho}$ (in the classical or quantum versions of the statistical mechanics, respectively) of the Canonical Ensembles consisting of absolutely isolated systems (classical or quantum), which should be representative of real systems in thermal contact with their surroundings at temperature T, can be directly postulated by appealing to their general forms (234) and (269), which in fact leave practically no freedom, given the fact that energies of such real systems can vary in very wide (infinite in the limit of infinitely large heat bath) ranges. It is nevertheless instructive to derive them by considering the system and the heat bath (representing its surrounding) as a single macroscopically isolated compound system and by applying to this supersystem the Microcanonical Ensemble.

We derive first the distribution of energy of the system in contact with the heat bath as this does not depend on whether it is treated classically or quantum mechanically. According to the rules of the probability theory, the probability that the total energy

---

[174] Actually as $3N-6$ oscillators because 6 generalized variables corresponding to translations and rotations of the solid as a whole are eliminated if it is kept at rest; however $3N-6$ is practically the same as $3N$.

[175] One can also remark that the notion of a "degree of freedom", fairly clear in mechanics, becomes, as one enters deeper into theoretical physics, more and more elusive and ultimately is used without any concrete content.

$E_{\rm tot}$ (fixed up to a small uncertainty ΔE) of the compound system is distributed as E and $E_{\rm tot}-E$ between the system and the heat bath is given by the relative number of (absolutely isolated) supersystems in the ensemble realizing such a partition of energy that is - because within the Microcanonical Ensemble all microstates (which correspond to the fixed, up to ΔE, energy $E_{\rm tot}$) are equally populated - by the ratio of the number of microstates (of the compound supersystem) corresponding to such a partition of the total energy $E_{\rm tot}$ to the total number of all microstates corresponding to this energy:

$$P(E,dE)=\frac{1}{\Gamma(E_{\rm tot},N_{\rm h.b.},N_{\rm sys},\dots,\Delta E)}\,\Delta E\,\omega_{\rm sys}(E)\,\omega_{\rm h.b.}(E_{\rm tot}-E)\,dE\,, \tag{289}$$

where

$$\Gamma(E_{\rm tot},N_{\rm h.b.},N_{\rm sys},\dots,\Delta E)=\Delta E\int_0^{\rm E_{tot}}dE\,\omega_{\rm sys}(E)\,\omega_{\rm h.b.}(E_{\rm tot}-E)\,. \tag{290}$$

Using the Microcanonical Ensemble definition $S_{\rm h.b.}(E)=k_{\rm B}\ln(\omega_{\rm h.b.}(E)\Delta E)$ of the entropy of the heat bath, the probability $P(E,dE)$ can be written as

$$P(E,dE)\propto dE\,\omega_{\rm sys}(E)\exp\left\{\frac{1}{k_{\rm B}}\,S_{\rm h.b}(E_{\rm tot}-E)\right\}.$$

Expanding now the expression in the exponent

$$S_{\rm h.b.}(E_{\rm tot}-E)=S_{\rm h.b.}(E_{\rm tot})-\left(\frac{\partial S_{\rm h.b.}(E_{\rm tot})}{\partial E_{\rm tot}}\right)_{N_{\rm h.b.},\dots}E+\dots, \tag{291}$$

and neglecting all the terms of higher order in E than the first one on account of the fact that in the limit of the very large heat bath the probability $P(E,dE)$ of distributions of energy markedly different from those in which $E/E_{\rm tot}\sim N_{\rm sys}/N_{\rm h.b.}$ should be negligible, while $S_{\rm h.b.}\sim N_{\rm h.b.}$ and $E_{\rm tot}\sim N_{\rm h.b.}$ (the n-th derivative of $S_{\rm h.b}$ with respect to $E_{\rm tot}$ is then $\sim N_{\rm h.b.}^{1-n}$ and the terms of higher order in E than the first one are suppressed with respect to it by powers of $N_{\rm sys}/N_{\rm h.b.}$), one can (in the limit of infinitely large heat bath) write the Canonical Ensemble distribution $\rho_E(E)$ of the system's energy in the form

$$\rho_E(E)={\rm Const.}\times\omega_{\rm sys}(E)\,e^{-E/k_{\rm B}T}\equiv{\rm Const.}\times\omega_{\rm sys}(E)\,e^{-\beta E}\,. \tag{292}$$

We have introduced here the traditional symbol

$$\beta\equiv\frac{1}{k_{\rm B}T}\,. \tag{293}$$

The normalization constant is of course given by[176]

$${\rm Const.}^{-1}=\int_0^\infty dE\,\omega_{\rm sys}(E)\,e^{-E/k_{\rm B}T}\,.$$

---

[176] To make it clear: the probability distribution defined by (289) and (290) is normalized by construction. However, once the expansion (291) has been truncated, the tails of the distribution are modified (insignificantly from the practical point of view) and its normalization must be readjusted.

Classical distribution function of the Canonical Ensemble
The Canonical Ensemble energy distribution (292) is independent of whether the system is treated classically or quantum mechanically. The corresponding phase space distribution function $\rho(q,p)$ of the classical Canonical Ensemble representative for the system in thermal equilibrium with its surrounding at temperature T can be obtained by marginalization (in the language of statistics) of the joint distribution function[177] $\rho(q,p,\,q_{\mathrm{h.b.}},p_{\mathrm{h.b.}})$ of the microcanonical ensemble representative for the real system and its surrounding treated as a single compound, macroscopically isolated supersystem:

$$\begin{aligned}\rho(q,p) &= \int_{E_{\mathrm{tot}}\leq\mathscr{H}_{\mathrm{h.b.}}+\mathscr{H}_{\mathrm{sys}}\leq E_{\mathrm{tot}}+\Delta E} d\Gamma_{(q_{\mathrm{h.b.}},\,p_{\mathrm{h.b.}})}\,\rho(q,p,\,q_{\mathrm{h.b.}},p_{\mathrm{h.b.}})\\ &\propto \int_{E_{\mathrm{tot}}-\mathscr{H}_{\mathrm{sys}}\leq\mathscr{H}_{\mathrm{h.b.}}\leq E_{\mathrm{tot}}+\Delta E-\mathscr{H}_{\mathrm{sys}}} d\Gamma_{(q_{\mathrm{h.b.}},\,p_{\mathrm{h.b.}})} = \exp\left\{\frac{1}{k_{\mathrm{B}}}\,S_{\mathrm{h.b.}}(E_{\mathrm{tot}}-\mathscr{H}_{\mathrm{sys}}(q,p))\right\}.\end{aligned}$$

It has been assumed here that the Hamiltonian of the compound (absolutely isolated) supersystems forming the ensemble is the sum $\mathscr{H}_{\mathrm{sys}}(q,p)+\mathscr{H}_{\mathrm{h.b.}}(q_{\mathrm{h.b.}},p_{\mathrm{h.b.}})$; in other words its part $\mathscr{H}_{\mathrm{sys-h.b.}}$ has been omited in agreement with the way the Microcanonical Ensemble considered here is constructed. Expanding $S_{\mathrm{h.b.}}(E_{\mathrm{tot}}-\mathscr{H}_{\mathrm{sys}}(q,p))$ in the power series in $\mathscr{H}_{\mathrm{sys}}(q,p)$, rejecting terms of this expansion of order higher than the first one (again, on account of the fact that they affect the constructed distribution function $\rho(q,p)$ very little - i.e. insignificantly from the practical point of view - in the limit of infinitely large heat bath) and including the first term of the expansion into the normalization constant Z_{stat}, one obtains the distribution function $\rho(q,p)$ of the classical Canonical Ensemble in the form

$$\rho(q,p)=\frac{1}{Z_{\mathrm{stat}}}\,\exp\{-\mathscr{H}_{\mathrm{sys}}(q,p)/k_{\mathrm{B}}T\}\,. \tag{294}$$

By construction it is the distribution function of the ensemble of absolutely isolated systems representative of a real system in thermal contact with its surrounding at temperature T. The normalization factor

$$Z_{\mathrm{stat}}(T,V,\dots,N)=\int d\Gamma_{(q,p)}\,\exp\{-\mathscr{H}_{\mathrm{sys}}(q,p)/k_{\mathrm{B}}T\}\,. \tag{295}$$

which is a function of the temperature T, the system's volume V, the number N (numbers $N_1,\dots,N_r$) of particles constituting it and possibly other macroscopic deformative parameters (denoted by the elipses) characterizing the system, is called the **Canonical**

---

[177] It is the distribution function of the Microcanonical Ensemble which, while being representative for the macroscopically isolated as a whole real supersystem consisting of the system and the heat bath exchanging between one another energy, is itself a set of isolated supersystems each of which consists of a system and a heat bath absolutely isolated not only from the rest of the world but also from one another. Still, the joint distribution function $\rho(q,p,q_{\mathrm{h.b.}},p_{\mathrm{h.b.}})$ conforms the equal probability principle, that is probabilities of all microstates corresponding to the total energy of the supersystem in the specified range are equal.

Statistical Sum or the **Canonical Partition Function**. By introducing under the integral in (295) the unity written as $1 = \int dE\, \delta(E - \mathscr{H}(q,p))$ and interchanging the order of the integrations, the canonical partition function can also be written in the alternative form as

$$Z_{\text{stat}}(T, V, \dots, N) = \int_0^\infty dE\, \omega(E)\, \exp\{-E/k_{\text{B}}T\}\,, \tag{296}$$

where $\omega(E)$ is the density defined (within the classical approach) in (239). As will be shown below, it contains the complete thermodynamic information about the system.

Statistical operator of the quantum Canonical Ensemble
The derivation of the statistical operator $\hat{\rho}$ of the quantum Canonical Ensemble representative for the system in thermal equilibrium with its surrounding at temperature T goes as follows. We consider the same setting as in the classical case. The basis of the Hilbert space $\mathcal{H} = \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{h.b.}}$ of the entire isolated compound supersystems (members of the considered Microcanonical Ensemble) is formed by state-vectors of the form $|l\rangle \otimes |l_{\text{h.b}}\rangle$ which are assumed to be eigenvectors of the Hamiltonian $\hat{\mathscr{H}}_{\text{sys}} \otimes \hat{1}_{\text{h.b.}} + \hat{1}_{\text{sys}} \otimes \mathscr{H}_{\text{h.b.}}$; the interaction term $\mathscr{H}_{\text{h.b.}-\text{sys}}$ ensuring the thermal contact between the real system and the real heat bath is, as in the classical case, absent.[178] Hence, energies of the two parts of the system are additive. The statistical operator of the corresponding Microcanonical Ensemble therefore reads (to make the notation easier we rename E_{tot} to E)

$$\begin{aligned}
\hat{\rho}_{\text{Micro}} &= \frac{1}{\Gamma(E \dots, \Delta E)} \sum_{\substack{l_{\text{h.b.}},\, l \\ E \le E_{l_{\text{h.b.}}} + E_l \le E + \Delta E}} (|l\rangle \otimes |l_{\text{h.b.}}\rangle)(\langle l_{\text{h.b.}}| \otimes \langle l|) \\
&= \frac{1}{\Gamma(E \dots, \Delta E)} \sum_{l} |l\rangle\langle l| \quad \otimes \sum_{\substack{l_{\text{h.b.}} \\ E - E_l \le E_{l_{\text{h.b.}}} \le E + \Delta E - E_l}} |l_{\text{h.b.}}\rangle\langle l_{\text{h.b.}}|\,.
\end{aligned}$$

The Canonical Ensemble statistical operator $\hat{\rho}$ is now obtained by taking the trace of $\hat{\rho}_{\text{Micro}}$ with respect to the Hilbert space $\mathcal{H}_{\text{h.b.}}$ of the heat bath.[179] The sum over $l_{\text{h.b.}}$ restricted to the heat bath states of energies between $E - E_l$ and $E + \Delta E - E_l$, where E_l is the energy of the system yields then just the factor $\Gamma_{\text{h.b.}}(E - E_l, N_{\text{h.b.}}, \dots, \Delta E) = \exp(S_{\text{h.b.}}(E - E_l)/k_{\text{B}})$. This, upon expanding up to the first order in E_l and including the zeroth order term in

---

[178] Let us stress again that the members of the considered Microcanonical Ensemble are the pairs consisting of the system and the heat bath absolutely isolated from one another (and from the surrounding).

[179] More precisely, its matrix elements $\langle l'|\hat{\rho}|l\rangle$ are identified with

$$\sum_{l_{\text{h.b.}}} \langle l_{\text{h.b.}}| \otimes \langle l'|\hat{\rho}_{\text{Micro}}|l\rangle \otimes |l_{\text{h.b.}}\rangle\,.$$

the normalization,[180] leads to the final result

$$\hat{\rho} = \frac{1}{Z_{\rm stat}} \sum_l |l\rangle\langle l| \exp(-E_l/k_{\rm B}T) \equiv \frac{1}{Z_{\rm stat}} \exp\Big(-\hat{\mathscr{H}}/k_{\rm B}T\Big)\,, \tag{297}$$

with the normalization factor (the quantum Canonical Ensemble partition function) $Z_{\rm stat}$ given by

$$\begin{aligned} Z_{\rm stat} &= {\rm Tr}\Big(e^{-\hat{\mathscr{H}}/k_{\rm B}T}\Big) = \sum_l \langle l|e^{-\hat{\mathscr{H}}/k_{\rm B}T}|l\rangle \\ &= \sum_l \exp(-E_l/k_{\rm B}T) = \sum_{E_n} d_n \exp(-E_n/k_{\rm B}T)\,. \end{aligned} \tag{298}$$

The first sum in the second line is over the system's Hamiltonian eigenvectors $|l\rangle$ and the second one is over the energy levels including their degeneracy factor d_n. When the system consists of interacting elements and its spectrum is quasi-continuous, the summation $\sum_{E_n} d_n$ is replaced by the integral $\int dE\,\omega(E)$ over the system's energy spectrum with the weight function $\omega(E)$ which is the density of quantum states corresponding to the system's energy between E and $E + dE$). This second form of the formula (298) is thus the quantum counterpart of the classical formula (296). It should be noted that both results (294) in the classical case and (297) have the forms (234) and (269), respectively, established by considering general properties of macroscopic systems.[181]

We will now show that if the system is simple in the thermodynamical sense, that is only a work associated with changing its volume can be done on it, the statistical sum given by (295) or (296) in the classical case and by (298) in the quantum case, is directly related to the Helmholtz free energy F which is already known to be the proper thermodynamic potential characterizing a system kept at a constant temperature by thermal contact with a heat bath. This readily follows from the "golden formulae" (285) or (284). Indeed, considering for the illustration the classical case, (285) applied to (294) yields:

$$S = -k_{\rm B}\overline{\ln\rho} = -k_{\rm B}\,\overline{(-\mathscr{H}/k_{\rm B}T - \ln Z_{\rm stat})} = \frac{1}{T}\,\overline{\mathscr{H}} + k_{\rm B}\ln Z_{\rm stat}\,.$$

After a small rearrangement and identifying $\overline{\mathscr{H}}$ with the system's internal energy U, this means that (recall, that $F = U - TS$)

$$F(T, V, N) = -k_{\rm B}T \ln Z_{\rm stat}(T, V, N)\,. \tag{299}$$

Obviously this formula stays valid in the quantum case (the steps are in that case completely analogous).

The rule (299) is also valid if the system consists of more than one material component - simply N should be replaced by the numbers $N_1, \dots, N_r$ of different kinds of particles.

---

[180] Justification of the rejection of all higher order terms of the expansion is the same as in the classical case.

[181] Notice also that the symbol β in the forms (234) and (269) has now acquired the interpretation (293).

Of course, if a multicomponent system consisting of r groups of identical (hence, quantum mechanically indistinguishable) particles is analysed classically, the measure over phase space is

$$d\Gamma_{(q,p)} = \prod_{a=1}^{r} \frac{d^{3N_a}q_a d^{3N_a}p_a}{N_a!\,(2\pi\hbar)^{3N_a}}\,.$$

As a result, the (classical) Canonical Ensemble formalism applied to a mixture of different classical perfect gases readily leads to the Gibbs Ansatz (167) and, as discussed in Lectures VIII and XI, it is the product of factors $N_1!\dots N_r!$ which leads to the mixing entropy.

Fluctuations of the system's energy

As already discussed, once the distribution function $\rho(q,p)$ in the classical case or the statistical operator $\hat{\rho}$ in the quantum case are given, the mean value of an observable O (represented by a phase space function $O(q,p)$ in te classical case and by a Hermitian operator $\hat{O}$ in the quatum case) can be computed as

$$\begin{aligned}\overline{O} &= \int d\Gamma_{(q,p)}\,O(q,p)\,\rho(q,p) = \frac{1}{Z_{\rm stat}}\int d\Gamma_{(q,p)}\,O(q,p)\,e^{-\mathscr{H}(q,p)/k_{\rm B}T}\,,\\ \overline{O} &= {\rm Tr}(\hat{O}\,\hat{\rho}) = \frac{1}{Z_{\rm stat}}{\rm Tr}\Big(\hat{O}\,e^{-\hat{\mathscr{H}}/k_{\rm B}T}\Big)\,,\end{aligned} \tag{300}$$

in the classical and quantum cases, respectively. Similarly can be computed also its mean quadratic fluctuation (228) to estimate its relative fluctuation (229) - recall (Lecture X) that the reliability of the results obtained using the approach based on the use of ensembles hinges on the smallness of relative fluctuations of the thermodynamical quantities calculated in this way. Since the energy of a system remaining in thermal equilibrium with a heat bath is not fixed, it is of interest to examine the possible magnitude of its fluctuations. As the Canonical Ensemble distribution function $\rho(q,p)$ or the statistical operator $\hat{\rho}$ are expressed through the system's Hamiltonian, the formulae for the mean system's energy $\overline{E} \equiv U$ and its fluctuation can be obtained in general terms, without using the explicit forms of these Hamiltonians. Indeed, from the formulae (300) it follows that (for definitess we consider the quantum case - classical calculation is completely analogous)

$$\begin{aligned}U = \overline{E} &= \frac{1}{Z_{\rm stat}}\,{\rm Tr}\Big(\hat{\mathscr{H}}\,e^{-\hat{\mathscr{H}}/k_{\rm B}T}\Big) = -\frac{1}{Z_{\rm stat}}\,\frac{\partial}{\partial\beta}\,{\rm Tr}\Big(e^{-\beta\hat{\mathscr{H}}}\Big)\Big|_{\beta=1/k_{\rm B}T}\\ &= -\,\frac{\partial}{\partial\beta}\,\ln Z_{\rm stat}\Big|_{\beta=1/k_{\rm B}T}\,.\end{aligned} \tag{301}$$

Using the formula (299) and noticing that $\partial/\partial\beta = -k_{\rm B}T^2\partial/\partial T$ this can be also written in the form

$$U = -T^2\left(\frac{\partial}{\partial T}\,\frac{F}{T}\right)_{V,N}\,,$$

which is just the standard thermodynamic identity (153). The mean quadratic fluctuation of the random variable E can be expressed in the similarly general way. To this end one starts with (we suppress the instruction to set $\beta = 1/k_{\rm B}T$ at the end)

$$\begin{aligned}\overline{E^2} &= \frac{1}{Z_{\rm stat}}\,{\rm Tr}\Big(\hat{\mathscr{H}}^2 e^{-\hat{\mathscr{H}}/k_{\rm B}T}\Big) = -\frac{1}{Z_{\rm stat}}\,\frac{\partial}{\partial\beta}\,{\rm Tr}\Big(\hat{\mathscr{H}}\, e^{-\beta\hat{\mathscr{H}}}\Big)\\ &= -\frac{1}{Z_{\rm stat}}\,\frac{\partial}{\partial\beta}\,\big(Z_{\rm stat}\,\overline{E}\big) = -\frac{\partial\overline{E}}{\partial\beta} - \overline{E}\,\frac{\partial}{\partial\beta}\,\ln Z_{\rm stat} \equiv -\frac{\partial\overline{E}}{\partial\beta} + \overline{E}^2\,.\end{aligned}$$

It follows that

$$\sigma_E^2 \equiv \overline{E^2} - \overline{E}^2 = -\frac{\partial\overline{E}}{\partial\beta} = k_{\rm B}T^2\left(\frac{\partial U}{\partial T}\right)_{V,N} = k_{\rm B}T^2 C_V\,. \tag{302}$$

Hence, in complete generality, the mean quadratic fluctuation of the system's energy is determined by its heat capacity (at constant volume) C_V. Since if the system is macroscopic, $C_V = Nc_v$ and $U = Nu$, the relative fluctuation of its energy is suppressed by $1/\sqrt{N}$:

$$\frac{\sqrt{\sigma_E^2}}{U} = \frac{1}{Nu}\,\sqrt{Nk_{\rm B}T^2c_V} \sim 1/\sqrt{N}\,,$$

in agreement with the general result (232). It should be also recalled that smallness of the fluctuations of thermodynamic quantities is a necessary conditions for equivalence of different ensembles.

It is instructive to obtain the same result by a different technique which at the same time will provide another justification - a one not relying on the "golden formulae" (285) or (284) - of the rule (299). To this end we rewrite the system's energy distribution (292), which has the same form, indeendently of wheter the classical or quantum Canonical Ensemble is used, in the form

$$\rho_E(E) = \frac{1}{Z_{\rm stat}\Delta E}\,\exp\{\beta\,(TS(E) - E)\}\,, \tag{303}$$

which exploits the Microcanonical Ensemble (Boltzmann) definition $S = k_{\rm B}\ln(\omega(E)\Delta E)$ of the entropy of the system (at fixed energy). If the system is large, $\rho_E(E)$ should be sharply peaked at $E = E^*$ which is practically equal to the mean $\overline{E} = U$. The exponent in the distribution (303) can be, therefore, expanded in the Taylor series around $E = E_* \approx U$ and the terms of this expansion higher than the second one can be discarded. The energy most probable value E_* is determined by the condition $d(TS(E) - E)/dE = 0$ which implies that the system's temperature $(\partial S(E)/\partial E)^{-1}$ is equal to the temperature T of the heat bath, and the truncated energy distribution takes the form

$$\rho_E(E) \approx \frac{1}{Z_{\rm stat}\Delta E}\,\exp\{\beta\,(TS(E_*) - E_*)\}\,\exp\left\{-\frac{1}{2k_{\rm B}T^2C_V}\,(E - E_*)^2\right\}. \tag{304}$$

If the normalization constant in the truncated form (304) is readjusted (and the domain of integration over energy is extended - in view of the negligible error this introduces - to the entire real axis; E_* becomes then identical with $\overline{E}$ given by the distribution obtained in this way), the system's energy fluctuation (302) is immediately recovered from the formula

$$\overline{(E-E_*)^2} = \text{Const.}\int_{-\infty}^{\infty} dE\,(E-E_*)^2 \exp\left\{-\frac{1}{2k_{\rm B}T^2C_V}\,(E-E_*)^2\right\}.$$

Furthermore, if one neglects the difference between the prefactor of (304) and the readjusted energy distribution normalization denoted "Const." in the formula above, one can write

$$Z_{\rm stat} = \frac{1}{\Delta E}\,e^{\beta(TS(E_*)-E_*)}\int_{-\infty}^{\infty} dE\,\exp\left\{-\frac{1}{2k_{\rm B}T^2C_V}\,(E-E_*)^2\right\},$$

which after evaluating the (gaussian) integral leads to the relation

$$\ln Z_{\rm stat} = -\beta\,(E_*-TS(E_*)) - \frac{1}{2}\ln\left(\frac{2\pi k_{\rm B}T^2C_V}{(\Delta E)^2}\right).$$

In the limit $N \gg 1$ the first term on the right hand side is of order N while the second one of order $\ln N$, i.e. can be neglected in comparison with the first one. Since E_* is in this regime identical with the system's mean energy U, the formula (299) is recovered.[182]

Canonical ensemble and magnetic systems

It is also worth to ask what thermodynamical potential is related to the Canonical Ensemble statistical sum $Z_{\rm stat}$ if the system is not simple but possesses e.g. magnetic properties. Let the (quantum) Hamiltonian of the system have the form (the subscript "int" stands here for "internal")

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_{\rm int} - \sum_{i=1}^{N}\hat{\boldsymbol{\mu}}_i\cdot\boldsymbol{\mathcal{H}}\,,$$

where the second term is the coupling of magnetic moments $\boldsymbol{\mu}_i$ of the system's individual elements (molecules, for instance) to an external magnetic field $\boldsymbol{\mathcal{H}}$ through the operators $\hat{\boldsymbol{\mu}}_i$ representing them.[183] Assuming such a system to be in equilibrium with a heat bath

---

[182] It may superficially seem that the "golden rule" gives the formula (299) without any approximations, while its derivation presented here strongly relies on the limit $N \gg 1$ and typicality of the system's energy spectrum. In connection with this it should be recalled that the justification of the "golden rule" also relies on similar assumptions so there is no real discrepancy here; the two derivations of the formula (299) simply differ by the stage at which one appeals to the essentially identical assumptions.

[183] In such a setting there is no distinction between the magnetic field $\boldsymbol{\mathcal{H}}_0$ produced by the experimental setup - by a current passing through a coil, for instance - in the absence of the magnetic material and the actual strength of the magnetic field $\boldsymbol{\mathcal{H}}$ when the magnetic material is present, because the magnetic moments $\boldsymbol{\mu}$ of the system's elements are implicitly assumed not to produce any magnetic field by themselves. Their mutual coupling, which in reality occurs mostly through the magnetic field they produce, is in statistical physics models usually taken into account in terms of their direct "contact" interactions of the form $-\sum_{i\neq j}J_{ij}\hat{\boldsymbol{\mu}}_j\cdot\hat{\boldsymbol{\mu}}_j$ (which, if present, are here included in $\hat{\mathscr{H}}_{\rm int}$).

at temperature T and applying the "golden formula" (284) to the Canonical Ensemble representing it we get

$$\begin{aligned} S &= -k_{\rm B}\overline{\left(-\frac{1}{k_{\rm B}T}\left(E_{\rm int}-\sum_i \boldsymbol{\mu}_i\cdot\boldsymbol{\mathcal{H}}\right)\right)} + k_{\rm B}\ln Z_{\rm stat} \\ &= \frac{1}{T}\,\overline{E_{\rm int}} - \frac{1}{T}\,\boldsymbol{\mathcal{H}}\cdot\sum_i \overline{\boldsymbol{\mu}_i} + k_{\rm B}\ln Z_{\rm stat}\,, \end{aligned}$$

Identifying now $\overline{E_{\rm int}}$ with the system's internal energy U and $\sum_i \overline{\boldsymbol{\mu}_i}$ with the system's total magnetization $\mathbf{M}$, one obtains

$$-k_{\rm B}T\ln Z_{\rm stat}(T,V,\boldsymbol{\mathcal{H}},N) = U - TS - \mathbf{M}\cdot\boldsymbol{\mathcal{H}}\,,$$

which is the magnetic Gibbs function $G(T,V,\boldsymbol{\mathcal{H}},N)$ the differential of which is[184]

$$dG = -SdT - p\,dV - \mathbf{M}\cdot d\boldsymbol{\mathcal{H}} + \mu\,dN\,.$$

The differential of the internal energy is then as usually (and always in this Course)

$$dU = TdS - p\,dV + \boldsymbol{\mathcal{H}}\cdot d\mathbf{M} + \mu\,dN\,.$$

The reader should be warned, however, that there is another school of authors who include the interaction of magnetic moments into what they call the system's internal energy (let us denote it $\tilde{U}$); the statistical sum is then related to the function $\tilde{F}$

$$-k_{\rm B}T\ln Z_{\rm stat}(T,V,\boldsymbol{\mathcal{H}},N) = \tilde{F} = \tilde{U} - TS\,,$$

This is however not only a mere change of the notation (using $\tilde{F}$ in place of G): since

$$\tilde{U} = U - \mathbf{M}\cdot\boldsymbol{\mathcal{H}}\,,$$

the differential of the "internal energy" $\tilde{U}$ is

$$d\tilde{U} = TdS - p\,dV - \mathbf{M}\cdot d\boldsymbol{\mathcal{H}} + \mu\,dN\,.$$

so that the elementary work of magnetization is now $-\mathbf{M}\cdot d\boldsymbol{\mathcal{H}}$ (if $\tilde{U}$ is used the Pippard's derivation of the elementary work must be modified). In any case the "golden formula" always allows to properly identify thermodynamical quantities related to the used ensemble.

Quantum Canonical Ensemble

Computing the partition function and the mean values of observables

$$Z_{\rm stat} = {\rm Tr}\Big(e^{-\hat{\mathscr{H}}/k_{\rm B}T}\Big)\,, \qquad \overline{O} = \frac{1}{Z_{\rm stat}}{\rm Tr}\Big(\hat{O}\,e^{-\hat{\mathscr{H}}/k_{\rm B}T}\Big)\,, \tag{305}$$

---

[184] The chemical potential μ should not be confused with the magnetic moment $\boldsymbol{\mu}$. It is clear that the SI system of units which brings in yet another μ_0 (the magnetic susceptibility of vaccuum) is utterly inconvenient here!

of the quantum Canonical Ensemble requires specifying the Hilbert space of the system of interest in which act the operators $\hat{\mathscr{H}}$, $\hat{\rho}$ and all operators $\hat{O}$ representing observables. If the system is, as usually is the case, composed of N elements, e.g. of N particles of some kind (we assume properties of these particles, like masses, spins, magnetic moments are known), quantum states of the world can be specified in terms of quantum states of individual elements (particles). The Hilbert space of the system can be then constructed as a tensor product of the Hilbert spaces spanned by state-vectors representing states of individual elements (particles), which for definiteness we will call single-particle states.[185] That is, one first specifies a basis of state-vectors representing quantum states of a single element taking for instance, if these elements are particles, the generalized eigenvectors $|\mathbf{x}\rangle$ of the position operator,[186] or the eigenvectors $|\mathbf{p}\rangle$ of the momentum operator or - if the particle has a nonzero spin - the vectors $|\mathbf{p},\sigma\rangle$, where σ is the spin projection onto a chosen axis (usually the z-axis). We will denote these basis vectors $|l\rangle$ (understanding the a single label l may stand for a set of independent labels). The vectors $|l_i\rangle$ thus span the single-particle Hilbert space $\mathcal{H}^{(1)}_i$ of the i-th particle (i-th element) of the system. The Hilbert space $\mathcal{H}^{(N)}$ of the entire system is then constructed as $\mathcal{H}^{(N)}=\mathcal{H}^{(1)}_1\otimes\ldots\otimes\mathcal{H}^{(1)}_N$, that is, it is spanned by the basis state vectors of the form

$$|l_1\rangle\otimes\ldots\otimes|l_N\rangle\,. \tag{306}$$

Most of physical systems however are composed of identical indistinguishable particles (or several groups of particles which are identical within these groups, but for definiteness we will consider only one type of identical particles - extension to several groups is more or less straightforward) and in this case the rules of quantum mechanics (which have been abstracted from experimental facts in the course of its developement) dictate that physical states of such systems are represented not by arbitrary superpositions of the state-vectors (306) but only by superpositions of totally antisymmetrized state-vectors (if the identical particles have half-integer spin, that is, are **fermions**) or superpositions of only totally symmetrized state-vectors (if the identical particles have integer spin, that is, are **bosons**). The roots of this rule are in special relativity - in four space-time dimensions states of half-integer spin particles must be antisymmetric while those of bosons must be symmetric if relativistically invariant quantum theories of their interactions, satisfying the so-called "local causality" requirements, are to be constructed (this is the celebrated **spin-statistics connection**); in two space dimensions other symmetry properties are also possible and the corresponding particles are called **anyons** (we will not consider this

---

[185] One sometimes meets with statements to the effect that the construction of the Hilbert space of a system as a tensor product of individual Hilbert spaces of its elements has something to do with the assumption that mutual interactions of these elements are (negligibly) weak. This is wrong. The choice of the Hilbert space in which quantum mechanics of a given system is realized is something which belongs to physics and goes beyond mathematics; one has first to decide - and this is what requires an amount of a physical insight - what are the possible quantum states (in the abstract sense) of the system - what states of its individual elements can physically be identified (be it only in principle) - and only then model them by the choice of the appropriate Hilbert space in which the action of operators is realized.

[186] These should better (and in fact can) be avoided, as the position operator does not exist in relativistic physics!

possibility here). Therefore the Hilbert spaces of a system of N fermions or of N bosons are spanned by the following basis vectors:

$$\begin{aligned}
&\frac{1}{\sqrt{N!}}\sum_P(-1)^P|l_{P(1)}\rangle\otimes\ldots\otimes|l_{P(N)}\rangle\equiv|l_1,l_2,\ldots,l_N\rangle \qquad \text{fermions}\,,\\
&\frac{1}{\sqrt{N!}}\sum_P|l_{P(1)}\rangle\otimes\ldots\otimes|l_{P(N)}\rangle\equiv|l_1,l_2,\ldots,l_N\rangle \qquad \text{bosons}\,.
\end{aligned}\tag{307}$$

The symbol P stands here for permutations of N labels and $(-1)^P$ denotes the permutation sign. These vectors will be denoted $|l_1,\ldots,l_N\rangle$. In both cases the labels l_i run over a countably infinite set of values; below we will assume that $l_i = 1,2,\ldots$ The number of labels l_i in each ket is, of course, equal N. It is convenient to adopt the convention (exploiting the antisymmetry or symmery of the vectors (307)) that the labels in kets are always ordered so that $l_1 < l_2 < \ldots < l_N$ in the case of fermions (no two labels l_i and l_j can in this case be equal - this is just the Pauli exclusion principle) and $l_1 \leq l_2 \leq \ldots \leq l_N$ in the case of bosons (many bosons can be simultaneously in the same single-particle state). Assuming that the single-particle state vectors $|l\rangle$ are normalized ($\langle l'|l\rangle = \delta_{l'l}$), the fermionic vectors (307) are automatically normalized to unity, while in the case of they require additional normalization if some labels l_i are equal.[187]

As the Hamiltonian $\hat{\mathscr{H}}^{(N)}$ of a system of N particles is supposed to be built (as appropriate tensor products) out of operators acting in the single-particle Hilbert spaces of individual particles, the traces in the formulae (305) can be computed using the bases (307):

$$\mathrm{Tr}\left(e^{-\hat{\mathscr{H}}/k_{\mathrm{B}}T}\right)=\frac{1}{N!}\sum_{l_1=1}^{\infty}\cdots\sum_{l_N=1}^{\infty}\langle l_N,\ldots,l_1|e^{-\hat{\mathscr{H}}/k_{\mathrm{B}}T}|l_1,\ldots,l_N\rangle\,,$$

and

$$\mathrm{Tr}(\hat\rho\,\hat O)=\frac{1}{N!}\sum_{l_1=1}^{\infty}\cdots\sum_{l_N=1}^{\infty}\frac{1}{N!}\sum_{l'_1=1}^{\infty}\cdots\sum_{l'_N=1}^{\infty}\langle l_N,\ldots,l_1|\hat\rho|l'_1,\ldots,l'_N\rangle\langle l'_N,\ldots,l'_1|\hat O|l_1,\ldots,l_N\rangle\,.$$

Notice that in these sums the orderings of labels l_i are not respected: The factors $1/N!$ cancel then multiple countings in these sums of the same state vectors written with different orderings of the labels l_i. This ensures proper taking into account of the normalization of the basis vectors. It should be stressed that in this form the formulae are completely general and include all possible quantum effects. In particular effects of the the mythical "quantum statistics". This will become clear as we come to consider the Grand Canonical Ensemble.

The adopted notation $|l_1,\ldots,l_N\rangle$ for the basis state-vectors is rather inconvenient, particularly when the considered particles are bosons. It is much more practical to pass

---

[187] If n_l bosons are in the same single-particle state $|l\rangle$, one has to multiply the state vector $|l_1,\ldots,l_N\rangle$ by the factor $1/\sqrt{n_l!}$ - see the formula (308).

to the so called **occupation number representation** in which the numbers in kets tell how many particles occupy successive single-particle states. Thus we set

$$|n_1, n_2, \dots\rangle = \frac{1}{\sqrt{n_1! n_2! \dots}} |1, \dots, 1, 2, \dots, 2, \dots\rangle . \tag{308}$$

The square root of the product of the factorials makes these vectors well normalized. Of course in the case of fermions $n_i = 0$ or 1 only. The sum of the occupation numbers n_i must always be N. For example, the (normalized to unity) vector

$$\frac{1}{\sqrt{3!2!2!}} |1, 1, 1, 2, 3, 3, 7, 7, 11, 13\rangle ,$$

of 10 bosons is of which three occupy the state $l = 1$, one is in the state $l = 2$, etc., in the occupation number representation written as

$$|3, 1, 2, 0, 0, 0, 2, 0, 0, 0, 1, 0, 1, 0, 0, 0, \dots\rangle .$$

It should be stressed that this is not a change of the basis in the Hilbert space but only a change of the notation. It is also important to note that in the notation $|l_1, \dots, l_N\rangle$ the number of entries in the ket is finite, but each label l_i can assume infinitely many (discrete values); in the notation $|n_1, n_2 \dots\rangle$, the number of entries in the ket is infinite but the values of the labels n_i are restricted by the condition $n_1 + n_2 + \dots = N$ (and, moreover, if particles are fermions each n_i is only either zero or one). The occupation number notation of state-vectors of N bosons should be also contrasted (in order to avoid a confusion) with the basis state-vectors $|n_1, \dots, n_N\rangle$ of a system of N (identical or not but distinguishable!) quantum harmonic oscillators: in this case it is the number of the labels n_i in each ket which is N but each n_i can run from 0 to infinity. We will see that when the restriction $n_1 + n_2 + \dots = N$ will get removed - and we will remove it passing to the Grand Canonical Ensemble - the system of bosons will become mathematically identical with a system of (infinitely many) harmonic oscillators.

Boltzmann approximation

The usefulness of the notation just explained will become evident when we introduce the creation and annihilation operators (associated with the single-particle states) through which all operators of interest acting in the system's Hilbert space - the Hamiltonian $\hat{\mathscr{H}}$ (including its interaction terms) of the system as well as all observables - can be expressed. Here, in order to illustrate the difficulty which exists, whether the particle interact with one another or not, we will consider only a system of N mutually noninteracting identical elements (atoms, molecules, spins) the joint Hamiltonian of which has the general form

$$\hat{\mathscr{H}}^{(N)} = \sum_{i=1}^{N} \hat{1}^{(1)} \otimes \dots \otimes \hat{\mathscr{H}}^{(1)}(i) \otimes \dots \otimes \hat{1}^{(1)} , \tag{309}$$

($\hat{\mathscr{H}}^{(1)}(i)$ at the i-th position) of a sum of N operators each of which acts essentially in only one single-particle Hilbert space. If the vectors $|l\rangle$ are chosen to be eigenvectors of the single particle Hamiltonian $\hat{\mathscr{H}}^{(1)}$ with with the eigenvalues ε_l, the action of the Hamiltonian (309) on the basis vectors $|n_1, n_2, n_3, \dots\rangle$ of the Hilbert space of N molecules (elements) is particularly simple

$$\hat{\mathscr{H}}^{(N)}|n_1, n_2, n_3, \dots\rangle = \left(\sum_l n_l\, \varepsilon_l\right) |n_1, n_2, n_3, \dots\rangle\,.$$

The statistical sum $Z_{\rm stat}$ of such a system of N mutually noninteracting molecules is given by the expression

$$Z_{\rm stat} = \sum_{n_1=0}^{n_{\rm max}} \sum_{n_2=0}^{n_{\rm max}} \dots \delta_{N, \sum_l n_l}\, e^{-n_1\varepsilon_1/k_{\rm B}T}\, e^{-n_2\varepsilon_2/k_{\rm B}T} \dots, \tag{310}$$

in which $n_{\rm max} = 1$, if the molecules are fermions and $n_{\rm max} = \infty$, if they are bosons. Unfortunately, even in this simple case the statistical sum cannot be computed easily (except for special forms of the spectra ε_l of the Hamiltonian of a single molecule) because of the presence of the Kronecker delta which expresses the condition of constancy of the number of the molecules in the Canonical Ensemble.

There are two ways of going around this diffuculty: one is to use the standard trick of statistical physics (and thermodynamics): in order to control o quantity which is absolutely conserved[188] by the system's dynamics we imagine that the system is in contact with a reservoir of this quantity allowing for exchanging it between the system and the reservoir (the Canonical Ensemble is itself an example of this trick: to control the system's internal energy we imagine it being exchanged with the heat bath at a fixed temperature and by controling the temperature of the heat bath we are able to control the mean energy of the system). This will lead to the Grand Canonical Ensemble which in general is representative for systems exchanging matter (molecules or other quantities) with their surroundings modeled by reservoirs of the matter kept at fixed chemical potentials μ. Another way (applicable only to systems composed of mutually noninteracting elements) is to use the so-called **Boltzmann approximation** which we now discuss.

---

[188] This is a subtle point: the quantity must be absolutely conserved not only by the interactions which are explicitly included in the Hamiltonian taken to perform actual computations but also by those interactions which are neglected on the account that they are weak but do exist in the real system and must be present for instance to let it reach the equilibrium. The example is the photon gas: photons interact with one another extremely weakly (linearity of the Maxwell equations means that "classical" photons do not interact with one another at all - their interaction is effectively induced by quantum effects) and this interaction which in principle can change their number can be safely neglected in computing statistical properties of the photon gas; moreover photons interact also with the walls of the container in which the gas is enclosed (the walls constantly absorb and emit photons, so their number is not conserved) and it is mostly these interactions, although they are neglected in computing statistical properties of the system, that are absolutely crucial for establishing the equilibrium. For those reasons one cannot apply this trick to the photon gas: the chemical potential of the gas of photons is zero. (Fortunately, the trick is not necessary - there are no restriction on the number of photons.)

In the Boltzmann approximation the expression (310) is replaced by

$$Z_{\rm stat} = \frac{1}{N!}\left(\sum_l e^{-\varepsilon_l/k_{\rm B}T}\right)^N . \tag{311}$$

In this approximation the statistical sum factorizes into the product of contributions of individual (mutually noninteracting) elements just as it does in analogous classical situations. The approximation (311) corresponds to the simple, essentially classical in its character, counting of the system's microstates: the number of ways of choosing n_1 particles which will be in the single-particle states $l = 1$ is

$$\binom{N}{n_1},$$

then the number of ways of choosing n_2 particles which will be in the single-particle state $l = 2$ out of the remaining $N - n_1$ particles is

$$\binom{N-n_1}{n_2},$$

and so on. This gives as the statistical weight (the degeneracy factor d_n in the formula (298)) of the level of energy $n_1\varepsilon_1 + n_2\varepsilon_2 + \dots$ the factor

$$\binom{N}{n_1}\binom{N-n_1}{n_2}\binom{N-n_1-n_2}{n_3}\dots = N!\left[\prod_{l=1}^{\infty} n_l!\right]^{-1} .$$

Notice that this counting treats particles as distinquishable (and does not take into account the Pauli exclusion principle); it corresponds to taking as the Hilbert space of the system of N identical elements the space spanned by all tensor products of the form (306) and not only by their totally symmetrized or totally antisymmetrized combinations. Indistinguishability of particles is here only taken into account by dividing all these factors by $N!$

With this counting the sum (310) is replaced by

$$Z_{\rm stat} = \frac{1}{N!}\sum_{n_1=0}^{N}\sum_{n_2=0}^{N}\dots\frac{N!}{\prod_{l=1}^{\infty}(n_l!)}\,\delta_{N,\sum_l n_l}\,e^{-n_1\varepsilon_1/k_{\rm B}T}\,e^{-n_2\varepsilon_2/k_{\rm B}T}\dots, \tag{312}$$

This precisely gives[189] the expression (311).

---

[189] The sum obtained in this way just the extension of the formula

$$(x_1+x_2)^N = \sum_{n=0}^{N}\binom{N}{n}x_1^n x_2^{N-n} = \sum_{n_1=0}^{N}\sum_{n_2=0}^{N}\frac{N!}{n_1!n_2!}\,\delta_{N,n_1+n_2}\,x_1^{n_1}x_2^{n_2} ,$$

to the expression $(x_1+x_2+x_3+\dots)^N$.

It is instructive to see on a simple example, what is the difference between the sums (310) in the cases when particles are bosons or fermions and the sum in (311). Let's take $N = 3$ particles and introduce the notation $x_l = \exp(-\varepsilon_l/k_{\rm B}T)$. If $N = 3$ the formula (310) yields

$$\begin{aligned} Z_{\rm stat} &= x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_1 x_2 x_3 + \dots \qquad && \text{bosons,} \\ Z_{\rm stat} &= x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 + \dots \qquad && \text{fermions,} \end{aligned}$$

while from the formula (311) we obtain (in both cases)

$$Z_{\rm stat} = \frac{1}{3!}\left(x_1 + x_2 + x_3 + x_4 + \dots\right)^3 = \frac{1}{6}\left(x_1^3 + 3x_1^2 x_2 + 3x_1 x_2^2 + 6x_1 x_2 x_3 + \dots\right).$$

It is clear that contributions to the statistical sum $Z_{\rm stat}$ of singly occupied single-particle states in (310) are properly taken into account in the approximation (311), but not those of multiply occupied single-particle states: if particles are fermions they are totally absent in (310) whereas if particles are bosons they have higher weights.

It follows that the Boltzmann approximation (311) should be reasonable in situations in which probabilities of multiple occupancy of the same single-particle states are low. This can be made quantitative only by going over to the Grand Canonical Ensemble in which the number of particles of the system is not fixed (this ensemble is representative for the system which can exchange energy and matter with a large - infinitely large in the limit - reservoir at temperature T and chemical potential μ). It is then possible to consider the mean number $\overline{n_l}$ of particles occupying a single-particle state $|l\rangle$. The formula which will be derived reads

$$\overline{n_l} = \frac{1}{\mp 1 + \exp((\varepsilon_l - \mu)/k_{\rm B}T)} \qquad \left\{\begin{array}{c}\text{bosons}\\ \text{fermions}\end{array}\right. .$$

This is smaller than 1 (that is, the probability of the occupation of a given energy state $|l\rangle$ is low) independently of the label l of the state, if the chemical potential (of the reservoir, and hence, also of the system which is in equilibrium with it) is large negative, i.e. when the **activity** z is small compared to unity:

$$z \equiv e^{\mu/k_{\rm B}T} \ll 1\,. \tag{313}$$

This is typically satisfied in rarefied perfect gases at temperatures not too low, so that the gas is not liquefied, and not too high so that the molecules do not dissociate yet and atoms are not ionized. The chemical potential of a perfect gas can be estimated by computing it within the classical Canonical Ensemble:

$$\mu = -k_{\rm B}T \ln\left[\frac{V}{N}\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2}\right]. \tag{314}$$

It becomes large negative at high temperatures (and moreover, $\mu/k_{\rm B}T \to -\infty$ as $T \to \infty$). More precisely, the condition (313) translates into the condition

$$\frac{V}{N}\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} \gg 1\,,$$

or

$$\lambda_T \equiv \left(\frac{2\pi\hbar^2}{mk_{\rm B}T}\right)^{1/2} \ll \left(\frac{V}{N}\right)^{1/3},$$

which means that the so-called **thermal wavelength** λ_T of the molecules should be small compared to the mean intermolecular distances.

It is now clear that the factor $1/N!$ included in the measure (243) over the classical phase space of a system of N identical particles results from the Boltzmann approximation to the proper quantum mechanical treatement of such particles.

The Boltzmann approximation to the statistical sum of a system of mutually nointeracting elements allows to develop the quantitative and fairly successful theory of specific heats of rarefied gases composed of molecules possessing internal structures which is valid at temperatures not too low (in order the Boltzmann approximation to be valid) and not too high (in order to allow neglecting processes of ionization and/or dissociation). The Boltzmann approximation allows then to neatly separate the motion of the centers of masses of the gas molecules from their internal excitations. To this end, the single-molecule Hilbert spaces $\mathcal{H}^{(1)}$ are taken to be spanned by the vectors of the form $|l\rangle = |\mathbf{p}, n\rangle \equiv |\mathbf{p}\rangle \otimes |n\rangle$ where $\mathbf{p}$ (if the molecules of the gas are enclosed in the volume $V = L^3$, the vectors $\mathbf{p}$ must be of the form $(2\pi/L)\mathbf{n}$, where $\mathbf{n}$ are the three-vectors of integer components) specifies the quantum state of the center of mass and n (which can be a multi-index) labels the internal energy levels of the single molecule. The action of the single molecule Hamiltonian on these vectors is

$$\hat{\mathscr{H}}^{(1)}|\mathbf{p}, n\rangle = \left(\frac{\mathbf{p}^2}{2m} + \varepsilon_n\right)|\mathbf{p}, n\rangle\,,$$

where m is the mass of the molecule and ε_n energies of its internal excitation. The factors $\sum_l \exp(-\varepsilon_l/k_{\rm B}T)$ in the formula (311) then read

$$\begin{aligned}\sum_l e^{-\varepsilon_l/k_{\rm B}T} &= \left(\sum_{\mathbf{p}} e^{-\mathbf{p}^2/2mk_{\rm B}T}\right)\left(\sum_n e^{-\varepsilon_n/k_{\rm B}T}\right)\\ &= \left(V\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\, e^{-\mathbf{p}^2/2mk_{\rm B}T}\right)\left(\sum_n e^{-\varepsilon_n/k_{\rm B}T}\right),\end{aligned}$$

(in the second line the sum over discrete wave vectors $\mathbf{p}$ has been converted into the integral using the standard prescription) and the entire statistical sum of the gas of mutually noninteracting molecules takes in the Boltzmann approximation the form

$$Z_{\rm stat} = \frac{V^N}{N!}\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3N/2}\left(\sum_n e^{-\varepsilon_n/k_{\rm B}T}\right)^N.$$

As a result the Helmholtza free nergy $F = -k_{\rm B}T\ln Z_{\rm stat}$ becomes the sum

$$F = -Nk_{\rm B}T\left\{1 + \ln\left[\frac{V}{N}\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2}\right] + \ln\left(\sum_n e^{-\varepsilon_n/k_{\rm B}T}\right)\right\}.$$

The effects of the molecules internal structure enters therefore this function additively and the same character have its contributions to internal energy, entropy and specific heats. In principle in the states $|n\rangle$ one should take into account all possible excitations of the nuclei of atoms out of which the molecules are buit as well as the quntum states of all their electrons. At temperatures $T \lesssim 10^5$ K most important are, however, only the lowest lying energy levels of electrons and the relative motion of the nuclei (their rotations and oscillations mentioned already in Lecture III) can be separated from the electronic one essentially by using the Feynman - Hellman theorem.

LECTURE XIII (STAT)

Grand Canonical Ensemble

As has been discussed, while the Canonical Ensemble applied to classical systems allows to effectively treat virtually all problems of systems consisting of mutually noninteracting elements (molecules), in the quantum case it is not as efficient because of the complicated structure of Hilbert spaces of states of indistinguishable elements composing physical systems: even if these indistinguishable elements (particles) are mutually noninteracting (in the sense of absence in the system's Hamiltonian of terms coupling them), there is an intrinsic quantum entanglement of their states resulting from the requirements of symmetry. As a result the quantum canonical statistical sum $Z_{\rm stat}$ cannot be computed (except for very special cases). Technically it is the condition of the constancy of the number of particles (elements) or, more generally, the strict conservation law of some physical quantity (like e.g. the electric charge) which makes the computation of $Z_{\rm stat}$ intractable practically. Therefore it is convenient to consider most of the problems of quantum statistical mechanics within the Grand Canonical Ensemble of systems which formally is representative for a system in equilibrium with its surrounding with which it can exchange heat (energy) and that quantity which is globally conserved.[190] The system's surrounding is modeled as a very big (in the limit infinitely large) reservoir at temperature T and chemical potential μ associated with the absolutely conserved quantity. As in most elementary applications this quantity is the number of (some kind of) particles, we will assume for definiteness, although this is not necessary, that the conserved quantity can take on discrete values and will denote it N.

The theoretically constructed ensemble representative of this real situation consists, as usually, of a set of $\mathcal{N}$ (the limit $\mathcal{N} \to \infty$ being implicitly understood) absolutely isolated systems which splits into groups of $\mathcal{N}_k$ (where $\sum_k \mathcal{N}_k = \mathcal{N}$, $\mathcal{N}_k \to \infty$ with the ratios $\mathcal{N}_k/\mathcal{N}$ kept fixed) systems with different values of the conserved quantity. The dynamics of each of these $\mathcal{N}$ systems forming the ensemble, which preserves the quantity of interest,[191] is set by the same Hamiltonian which is the Hamiltonian of the real system but with the terms coupling it to the surrounding - the reservoir - (and, perhaps, some other terms which are believed to be irrelevant from the practical point of view) removed. Although the numbers of particles (the values of the absolutely conserved quantity) and the internal energies of the different isolated and closed systems of the ensemble are then different, in all situations in which the computed within this ensemble fluctuations of the number of particles (of the quantity conserved in the real supersystem consisting of the system and the reservoir) around the mean value $\overline{N}$ are negligible (essentially proportional

---

[190] Extension of the formalism which will be developed to cases in which several quantities are absolutely conserved is straightforward.

[191] As already explained at the end of Lecture XII, the possibility of applying the Grand Canonical Ensemble technique relies on the *strict* conservation of the quantity of interest by the complete Hamiltonian of the real system and its surrounding and not only by those terms of the system's Hamiltonian which have been retained (in the course of simplifying the real problem to a tractable form) for doing actual statistical computations exploiting the ensemble of systems constructed theoretically.

to $1/\overline{N}^{1/2}$), all thermodynamical results obtained using the Grand Canonical Ensemble are from the practical point of view the same as the ones that would be obtained using the Canonical Ensemble representing the real system (in thermal contact with a heat bath at temperature T) with the fixed number $N = \overline{N}$ of particles. Similarly, if the computed within this ensemble fluctuations of the number of particles and of energy around their means $\overline{N}$ and $\overline{E}$ are small, the results can be applied to real macroscopically isolated and closed systems (for which the proper representative ensemble would otherwise be the microcanonical ensemble) with fixed energy $U = \overline{E}$ and fixed number $N = \overline{N}$ of particles.

As usually, the central problem is to theoretically construct the ensemble that is, to choose in the classical case the set of distribution functions $\rho(q, p, N)$ (specifying the density of the points in the phase spaces of N-particle systems and the distribution of systems with different values of N in the ensemble) or in the quantum case the appropriate statistical operator $\hat{\rho}$ (this will require introducing first the appropriate Hilbert space). This can be done by considering the system and its surrounding (the reservoir of heat and of N) as a macroscopically isolated supersystem and by associating with it the microcanonical ensemble. We will derive in this way first the probability distribution of energy and of the number of particles in the system of interest, because this (joint) distribution is independent of wheter the problem is treated within classical or quantum statistical mechanics.

Grand Canonical Ensemble joint energy-particle number distribution

Thus we assume that the real supersystem consists of a big reservoir weakly interacting with the system of interest and that particles (molecules) which are treated as indestructible (or the quantity which is conserved) can be exchanged between the system and the reservoir. The entire supersystem is macroscopically isolated and characterized by the total energy E_{tot} (with some uncertainty ΔE) and the total number of particles N_{tot} (total value of the conserved quantity; if this can assume continuous values one has to also assume that it too is fixed up to some uncertainty). The corresponding microcanonical ensemble representative for the real physical supersystem consists of $\mathcal{N}$ ($\mathcal{N} \to \infty$ in the limit) absolutely isolated supersystems composed of the reservoir and the system; it is however essential that in this ensemble the particles (the conserved quantity) and energy are not allowed to be exchanged between the reservoir and the system (the reservoir and the system are separated by completely isolating walls): the total energy E_{tot} of the supersystem and its total number N_{tot} of particles are distributed between the system and the reservoir in all possible ways and - in agreement with the principle of a priori equal probabilities - all microstates of the supersystem characterized by a concrete distributions of E_{tot} and N_{tot} have the same statistical weight. According to the rules of the probability theory the probability $\rho_E(N, E)dE$ that there are N particles in the system (and $N_{\text{tot}} - N$ particles in the reservoir) and that its energy is between E and $E + dE$ is given by the relative number of supersystems in the ensemble realizing this partition that is, because of a priori equal probabilities, by the ratio of the number of microstates realizing this situation to the total number of all microstates of the entire supersystem (the subscripts "res" and "sys" refer to the reservoir and the system, respectively)

$$\rho_E(N,E)\,dE = \text{Const.}\,\omega_{\text{sys}}(E,N)\,\omega_{\text{res}}(E_{\text{tot}}-E,\,N_{\text{tot}}-N)\,\Delta E\,dE\,,$$

where

$$\text{Const.}^{-1} = \sum_{N=0}^{N_{\text{tot}}}\int_0^{E_{\text{tot}}} dE\,\omega_{\text{sys}}(E,N)\,\omega_{\text{res}}(E_{\text{tot}}-E,\,N_{\text{tot}}-N)\,\Delta E\,.$$

The quantity $\omega_{\text{res}}(E_{\text{tot}}-E,\,N_{\text{tot}}-N)\,\Delta E$ is however directly related to the entropy of the reservoir treated as a macroscopically isolated system with energy $E_{\text{tot}}-E$ and the number $N_{\text{tot}}-N$ of particles. Thus

$$\rho_E(E,N) \propto dE\,\omega_{\text{sys}}(E,N)\,\exp\left\{\frac{1}{k_{\text{B}}}\,S_{\text{res}}(E_{\text{tot}}-E,\,N_{\text{tot}}-N)\right\},$$

and, upon expanding this entropy to the first order in E and N (the justification of truncating the expansion being essentially the same as in the case of the canonical ensemble), taking the limit of infinitely large reservoir (and therefore, implicitly letting E_{tot} and N_{tot} to tend to infinity with the ratio $E_{\text{tot}}/N_{\text{tot}}$ kept constant - in this way the constructed distribution $\rho_E(E,N)$ extends to an arbitrarily big number N of particles and arbitrarily large energies E of the system) and normalizing the distribution anew (to correct for the inessential change in its tails introduced by dropping the higher order terms of the expansion), one obtains the distribution

$$\rho_E(E,N) = \text{Const.}\ \omega_{\text{sys}}(E,N)\ \exp\left\{-\frac{1}{k_{\text{B}}T}\,E + \frac{\mu}{k_{\text{B}}T}\,N\right\}, \tag{315}$$

in which T and μ are to be identified with the temperature and the chemical potential of the reservoir and

$$\text{Const.}^{-1} = \sum_{N=0}^{\infty}\int_0^{\infty} dE\,\omega_{\text{sys}}(E,N)\ \exp\left\{-\frac{1}{k_{\text{B}}T}\,E + \frac{\mu}{k_{\text{B}}T}\,N\right\},$$

The mean number of particles in the system and/or its mean energy as well as the fluctuations of these quantities and their correlations can now be computed using directly the distribution (315).

Classical phase space distribution

The corresponding classical distribution functions $\rho(q,p,N)$ of the Grand Canonical Ensemble representative for the system in thermal and material contact with its surrounding (necessary for instance to compute mean values of quantities other than energy and particle number) is derived by marginalizing (in the statistical sense) the distributions ρ^N_{micro} corresponding to N particles in the system (and $N_{\text{tot}}-N$ in the reservoir) of the supersystem's microcanonical ensemble

$$\rho^N_{\text{micro}} = \begin{cases} \text{Const.}, & \text{if}\quad E_{\text{tot}} \le \mathscr{H}^{(N)}_{\text{sys}} + \mathscr{H}^{(N_{\text{tot}}-N)}_{\text{res}} \le E_{\text{tot}} + \Delta E \\ 0\,, & \text{otherwise} \end{cases}$$

in which Const.$^{-1} = \Gamma(E_{\rm tot}, V_{\rm res}, V_{\rm sys}, N_{\rm tot}, \Delta E)$ is given by

$$\sum_{N=0}^{N_{\rm tot}} \int d\Gamma^{(N)}_{\rm sys} \int d\Gamma^{(N_{\rm tot}-N)}_{\rm res} \theta(E_{\rm tot} + \Delta E - \mathscr{H}^{(N)}_{\rm sys} - \mathscr{H}^{(N_{\rm tot}-N)}_{\rm res})\, \theta(\mathscr{H}^{(N)}_{\rm sys} + \mathscr{H}^{(N_{\rm tot}-N)}_{\rm res} - E_{\rm tot})\,,$$

with respect to the variables of the reservoir. That is,

$$\begin{aligned} \rho(q,p,N) &= \text{Const.} \int_{E_{\rm tot}-\mathscr{H}^{(N)}_{\rm sys} \le \mathscr{H}^{(N_{\rm tot}-N)}_{\rm res} \le E_{\rm tot}+\Delta E-\mathscr{H}^{(N)}_{\rm sys}} d\Gamma^{(N_{\rm tot}-N)}_{\rm res} \\ &= \text{Const.} \times \Gamma_{\rm res}(E_{\rm tot} - \mathscr{H}^{(N)}_{\rm sys}, V_{\rm res}, N_{\rm tot} - N, \Delta E)\,. \end{aligned}$$

This again is proportional to the exponential of the reservoir's entropy $(1/k_{\rm B})S_{\rm res}(E_{\rm tot} - \mathscr{H}^{(N)}_{\rm sys}, V_{\rm res}, N_{\rm tot} - N)$, so upon expanding to the first order, taking the limits $E_{\rm tot} \to \infty$, $N_{\rm tot} \to \infty$ with the ratio $E_{\rm tot}/N_{\rm tot}$ kept constant and re-adjusting the normalization one arrives at the final result

$$\rho(q,p,N) = \frac{1}{\Xi_{\rm stat}} \exp\left\{-\frac{1}{k_{\rm B}T}\left(\mathscr{H}^{(N)}(q,p) - \mu\, N\right)\right\}, \tag{316}$$

in which the Grand Canonical Ensemble statistical sum $\Xi_{\rm stat}(T,V,\mu)$ (in case only the volume work can be done on the system in question) is given by

$$\begin{aligned} \Xi_{\rm stat} &= \sum_{N=0}^{\infty} \int d\Gamma^{(N)}_{(q,p)} \exp\left\{-\frac{1}{k_{\rm B}T}\left(\mathscr{H}^{(N)}(q,p) - \mu\, N\right)\right\} \qquad (317) \\ &= \sum_{N=0}^{\infty} e^{\mu N/k_{\rm B}T} \int d\Gamma^{(N)}_{(q,p)} \exp\left(-\mathscr{H}^{(N)}(q,p)/k_{\rm B}T\right) \equiv \sum_{N=0}^{\infty} e^{\mu N/k_{\rm B}T}\, Z_{\rm stat}(T,V,N)\,. \end{aligned}$$

The quantity $z \equiv \exp(\mu/k_{\rm B}T)$ is frequently called **activity**. Using the formula (296) the statistical sum $\Xi_{\rm stat}$ of the Grand Canonical Ensemble can be also written in the form

$$\Xi_{\rm stat} = \sum_{N=0}^{\infty} e^{\beta\mu N} \int dE\, \omega(E,N)\, e^{-\beta E}\,, \tag{318}$$

in which $\omega(E,N)$ is the N-particle system energy density. The last expression for $\Xi_{\rm stat}$ in (317) as well as the formula (318) will be valid also within the quantum version of the grand canonical ensemble. Moreover, the comparison with the system's joint energy-particle number distribution (315) shows that the normalization factor denoted "Const." in the latter formula is precisely equal $1/\Xi_{\rm stat}$.

One can now use the entropy golden formula (285) to relate the statistical sum $\Xi_{\rm stat}$ to the appropropriate thermodynamical potential:

$$\begin{aligned} S = -k_{\rm B}\, \overline{\ln\rho} &= -k_{\rm B} \sum_{N=0}^{\infty} \int d\Gamma^{(N)}\, \rho(q,p,N) \ln\rho(q,p,N) \\ &= \frac{1}{T} \sum_{N=0}^{\infty} \int d\Gamma^{(N)}\, \rho(q,p,N) \left(\mathscr{H}^{(N)}(q,p) - \mu\, N\right) + k_{\rm B} \ln\Xi_{\rm stat}\,. \end{aligned}$$

Identifying the first term on the right hand side with $\overline{E}/T$, the second one with $\mu\overline{N}/T$ and rearranging the formula one obtains (we denote $\overline{E}$ by U and drop the bar on N)

$$-k_{\rm B}T\ln\Xi_{\rm stat}(T,V,\mu)=U-TS-\mu\,N\,. \tag{319}$$

Thus, the Grand Canonical Ensemble partition function is related to the thermodynamical potential (151)

$$\Omega(T,V,\mu)=U-TS-\mu N=-V\,p(T,\mu)\,. \tag{320}$$

(The last form follows from extensiveness - see Lectures VII and VIII.) It is given as a function of its natural variables and contains, therefore, the complete information about the system's thermodynamics. Of course, if additional reversible works can be done on the system, for instance if the system has magnetic properties and is placed in an external magnetic field $\boldsymbol{\mathcal{H}}$, the function $-k_{\rm B}T\ln\Xi_{\rm stat}(T,V,\boldsymbol{\mathcal{H}},\mu)$ is the appropriate thermodynamical potential $\Phi=U-TS-\boldsymbol{\mathcal{H}}\cdot\mathbf{M}-\mu\,N$ (provided we continue not to include the interaction of the system magnetization with the external field into internal energy).

Fluctuations of the number of particles

Before we embark ourselves in developing the mathematical formalism necessary to discuss the quantum version of the Grand Canonical Ensemble, let us estimate the magnitude of fluctuations of the number of particles in a simple system (a fluid, for instance). Since the distribution (315) which is sufficient to compute it is independent of whether one works within classical or quantum ensembles (it must follow within both approaches), one can use for this purpose the formula (317) obtained so far using the classical version of the ensemble. Using this formula the mean number of particles in the system can be computed as ($\beta\equiv1/k_{\rm B}T$)

$$\overline{N}=\frac{1}{\Xi_{\rm stat}}\sum_{N=1}^{\infty}N\,e^{\beta\mu N}\,Z_{\rm stat}(T,V,N)=\frac{1}{\beta\,\Xi_{\rm stat}}\frac{\partial}{\partial\mu}\sum_{N=1}^{\infty}e^{\beta\mu N}\,Z_{\rm stat}(T,V,N)=\frac{1}{\beta}\frac{\partial}{\partial\mu}\ln\Xi_{\rm stat}\,.$$

To compute the mean quadratic fluctuation $\sigma_N^2=\overline{(N-\overline{N})^2}=\overline{N^2}-\overline{N}^2$ one uses the standard trick and writes

$$\begin{aligned}\overline{N^2}&=\frac{1}{\Xi_{\rm stat}}\sum_{N=1}^{\infty}N^2\,e^{\beta\mu N}\,Z_{\rm stat}(T,V,N)=\frac{1}{\beta\,\Xi_{\rm stat}}\frac{\partial}{\partial\mu}\sum_{N=1}^{\infty}N\,e^{\beta\mu N}\,Z_{\rm stat}(T,V,N)\\&=\frac{1}{\beta\,\Xi_{\rm stat}}\frac{\partial}{\partial\mu}\left(\Xi_{\rm stat}\,\overline{N}\right)=\frac{1}{\beta}\frac{\partial\overline{N}}{\partial\mu}+\frac{\overline{N}}{\beta}\frac{\partial}{\partial\mu}\ln\Xi_{\rm stat}\,.\end{aligned}$$

(In the penultimate step the factor $\Xi_{\rm stat}/\Xi_{\rm stat}$ has been inserted between the derivative and the summation sign). Since the last term is just $\overline{N}^2$, one arrives at the result

$$\sigma_N^2=k_{\rm B}T\left(\frac{\partial N}{\partial\mu}\right)_{V,T} \tag{321}$$

In the last formula the mean $\overline{N}$ has been identified with the thermodynamical variable N. To bring this result to a more informative form, one imagines that $N = N(T,\, V,\, p(T,\mu))$ to write it as

$$\sigma_N^2 = k_{\rm B}T\left(\frac{\partial N}{\partial p}\right)_{V,T}\left(\frac{\partial p}{\partial \mu}\right)_T .$$

One can now use the Gibbs-Duhem differential relation (161) but now with s and v interpreted as quantities per particle

$$d\mu = -s\,dT + v\,dp\,,$$

which implies that at constant T the derivative of p with respect to μ is $1/v$. Furthermore, since the first derivative is taken at constant volume, one can write it as the derivative of $1/v \equiv N/V$, instead of the derivative of N:

$$\sigma_N^2 = k_{\rm B}TV\left(\frac{\partial}{\partial p}\frac{1}{v}\right)_T\frac{1}{v} = -k_{\rm B}T\,\frac{V}{v^3}\left(\frac{\partial v}{\partial p}\right)_T = \frac{k_{\rm B}T}{v^2}\,k_T\,V = \frac{N}{v}\,k_{\rm B}T\,k_T\,. \tag{322}$$

Thus, the mean quadratic fluctuation of the number of particles in the ensemble is proportional to the system's isothermal compressibility k_T and the volume V. Hence, in normal circumstances $\sigma_N^2 \propto N$ (because $V \propto N$) and the relative fluctuation $\sqrt{\sigma_N^2}/N$ is tiny being of order $1/\sqrt{N}$, in agreement with the general considerations of Lecture X based on the principle of statistical independence. The number of particles in the system is therefore, practically equal to the mean and the results obtained by applying the Grand Canonical Ensemble to the system are practically the same as would be the ones obtained using the Canonical Ensemble in which the number of particles is strictly fixed. The two ensembles are from the practical point of view equivalent as far as the predictions are concerned (but may be not equivalent as far as the computational difficulties are concerned - this is why using the Grand Canonical Ensemble may be preferred). Thus, going over to the Grand Canonical Ensemble is yet another application of the same general idea: to keep fixed in the system the value of a quantity which is strictly conserved by the dynamics (not only by the interactions which are explicitly taken into account in statistical mechanics computations but also by those which are so weak that can be neglected form the practical point of view but in fact ensure establishing equilibrium), one imagines it being exchanged between the system and an (infinitely) large reservoir at an appropriate chemical potential of this quantity. By controling this chemical potential one controls the mean value of the quantity of interest in the system. If the fluctuations of this quantity (quantified by its mean quadratic fluctuation and, perhaps, other moments of its probability distribution) in this generalized setting are small, the results should agree with those obtained withn the approach in which the value of this quantity in the system is strictly fixed.

The two approaches (the one using the canonical ensemble and the one using the grand ensemble) cesse however to give the same result in some special situations: near the critical point of a fluid k_T becomes very large and diverges at the critical point; the fluctuations of the number of particles in the system (think of a small subvolume of a fluid enclosed in a large container as the system) then strongly fluctuates and this is reflected in strong fluid density fluctuations. This is responsible for the phenomenon of the so-called critical opalescence - the fluctuations of the fluid occur on all length scales and all wavelength of light directed onto the fluid are strongly scattered by the density fluctuations.

It is instructive to compute the particle number fluctuations using the distribution (315) integrated over energy:

$$\begin{aligned} P_N(N) &= \int dE\, \rho_E(E,N) = \frac{1}{\Xi_{\rm stat}}\, e^{\beta\mu N}\, Z_{\rm stat}(T,V,N) \\ &= \frac{1}{\Xi_{\rm stat}}\, \exp\{-\beta(F(T,V,N)-\mu N)\}\,. \end{aligned}$$

Since in normal conditions (i.e. not close to a critical point) the distribution $P_N(N)$ should be extremely sharply peaked at N_* which is practically equal to $\overline{N}$, one can expand the exponent in the taylor series around N_* defined by the condition $d(F-\mu N)/dN = 0$. The second derivative of F which gives the coefficient of the first nontrivial term of this series (the zeroth order term combines with the normalization factor of $P_N(N)$ and the first order term vanishes) can be computed by writing $F(T,V,N) = Nf(T,v) \equiv Nf(T,V/N)$:

$$\frac{\partial^2 F}{\partial N^2} = 2\,\frac{\partial f(T,v)}{\partial N} + N\,\frac{\partial^2 f(T,v)}{\partial N^2} = \frac{v^2}{N}\,\frac{\partial^2 f}{\partial v^2} = \frac{v}{N}\,\frac{1}{k_T}\,.$$

After the expansion of the exponent in the distribution $P_N(N)$ is truncated at the second term, the summation over N replaced by integration over the entire real N-axis (in view of the shape of the distribution this should introduce only a negligibly small error) and the normalization factor readjusted, the result (322) is readily recovered.

On the other hand if (in the limit of a large system) the difference between the readjusted normalization factor of the truncated distribution $P_N(N)$ and the original one, $1/\Xi_{\rm stat}$ is disregarded, one can write

$$\Xi_{\rm stat} = e^{-\beta(F(T,V,N_*)-\mu N_*)} \int dN\, \exp\left\{-\frac{1}{2\sigma_N^2}\,(N-N_*)^2\right\},$$

which, after the (gaussian) integral is evaluated, gives the relation

$$\ln \Xi_{\rm stat} = -\beta(F(T,V,N_*)-\mu N_*) - \frac{1}{2}\ln(2\pi\sigma_N^2)\,.$$

Since the first bracket on the right hand side is of order N_*, while the last term - of order $\ln N_*$, in the thermodynamic limit one obtains in this way another justification (which does not use the “golden formula” for entropy) of the formula (319) central to all thermodynamical applications of the Grand Canonical Ensemble.

Quantum Grand Canoncal Ensemble

As has been explained, it is the quantum case in which the transition to the Grand Canonical Ensemble is of great help even in the situations in which the number of particles (molecules) in the system is really fixed because the Canonical Ensemble statistical sum is intractable exactly even in the simplest case of mutually noninteracting particles (elements). One then imagines that the system is coupled with a reservoir with which it exchanges energy and particles and constructs a statistical ensemble corresponding to this situation. If the fluctuations of the number of particles in the system around the mean value $\overline{N}$ are negligibly small (typically suppressed by the factor $1/\overline{N}^{1/2}$) all thermodynamical results derived using the Grand Canonical Ensemble are practically the same as those which would be obtained if the Canonical Ensemble partition function $Z_{\rm stat}$ could be computed exactly.

The microcanonical ensemble corresponding to the system of interest exchanging matter and energy with a reservoir consists of a set of $\mathcal{N}$ (with the limit $\mathcal{N}\to\infty$ implicitly understood) **absolutely isolated** supersystems each consisting of a copy of the system and of the reservoir which are both microscopically **isolated and closed** i.e. **do not** exchange between one another particles or energy. Still the distribution of the supersystem's total energy $E_{\rm tot}$ and of its total number $N_{\rm tot}$ of particles between the system and the reservoir follows the principle of a priori equal weights (because it is intended to reflect the situation of the real system which does exchange energy and matter with the reservoir). Owing to this, for each particular partition of $N_{\rm tot}$ into $N_{\rm r}$ and N it is possible to introduce two separate Hilbert spaces $\mathcal{H}^{(N)}_{\rm sys}$ and $\mathcal{H}^{(N_{\rm r})}_{\rm res}$ spanned respectively by the vectors[192] $|l_1,\dots,l_N\rangle$ and $|L_1,\dots,L_{N_{\rm r}}\rangle$ defined in (307). In each of the Hilbert spaces $\mathcal{H}^{(N)}_{\rm sys}$ of the system (Hilbert spaces $\mathcal{H}^{(N_{\rm r})}_{\rm res}$ of the reservoir) acts the corresponding Hamiltonian $\hat{\mathscr{H}}^{(N)}_{\rm sys}$ of the system consisting of N particles (the corresponding Hamiltonian $\hat{\mathscr{H}}^{(N_{\rm r})}_{\rm res}$ of the reservoir consisting of $N_{\rm r}=N_{\rm tot}-N$ particles). We assume here that each of the Hamiltonians, $\hat{\mathscr{H}}^{(N)}_{\rm sys}$ of the N-particle system and $\hat{\mathscr{H}}^{(N_r)}_{\rm res}$ of the N_r-particle reservoir can be split into an interaction term and a free part which has the form of the sum (309) of the single particle Hamiltonians and that the vectors $|l_1,\dots,l_N\rangle$ and $|L_1,\dots,L_{N_{\rm r}}\rangle$ are the eigenvectors of the free parts of the respective Hamiltonians of the system and of the reservoir.[193] It is clear that because the real system is viewed as exchanging particles with the reservoir, both types of vectors must be either symmetric (if exchanged particles are bosons) or antisymmetric (if the exchanged particles are fermions) in their labels $l_1,\dots,l_N$ and $L_1,\dots,L_{N_{\rm r}}$. These vectors can also be written in the occupation number representation discussed in Lecture XII. We **do not** assume here that the

---

[192] If particles are bosons, additional factors ensuring their proper normalization should be included in those vectors in which some of the labels in the set $l_1,\dots,l_N$ (in the set $L_1,\dots,L_{N_{\rm r}}$) assume the same values.

[193] Since this is to model the real supersystem in which one kind of particles are exchanged between the system and the reservoir, it is clear that the Hamiltonians $\hat{\mathscr{H}}^{(N)}_{\rm sys}$ and $\hat{\mathscr{H}}^{(N_r)}_{\rm res}$ are built out of the same the single-particle Hamiltonians $\hat{\mathscr{H}}^{(1)}$ and the structures of the Hilbert spaces $\mathcal{H}^{(N)}_{\rm sys}$ and $\mathcal{H}^{(N_{\rm r})}_{\rm res}$ are the same - they differ only in the number of copies of the single-particle spaces $\mathcal{H}^{(1)}$. We do not however need to assume that the interaction terms of the Hamiltonians $\hat{\mathscr{H}}^{(N)}_{\rm sys}$ and $\hat{\mathscr{H}}^{(N_r)}_{\rm res}$ are of the same sort.

Hamiltonians $\hat{\mathscr{H}}^{(N)}_{\rm sys}$ and $\hat{\mathscr{H}}^{(N_r)}_{\rm res}$ correspond to nointeracting systems. Still, in each of the separate Hilbert spaces $\mathcal{H}^{\rm sys}_{(N)}$ of the system ($\mathcal{H}^{\rm res}_{(N_{\rm r})}$ of the reservoir) we can, as is always possible when quantum systems are not of infinite spatial extent, introduce a basis $|n^{(N)}\rangle$ ($|n^{(N_{\rm r})}_{\rm res}\rangle$) formed by proper (i.e. normalizable) eigenvectors of the Hamiltonian $\hat{\mathscr{H}}^{(N)}_{\rm sys}$ (of the Hamiltonian $\hat{\mathscr{H}}^{(N_{\rm r})}_{\rm res}$). If these Hamiltonians are not trivial, the eigenvectors $|n^{(N)}\rangle$ ($|n^{(N_{\rm r})}_{\rm res}\rangle$) are built as, perhaps very complicated, linear combinations of the basis vectors $|l_1,\dots,l_N\rangle$ ($|L_1,\dots,L_{N_{\rm r}}\rangle$). Only if the elements (particles) of the system (reservoir) are mutually noninteracting are the Hamiltonian eigenvectors $|n^{(N)}\rangle$ ($|n^{(N_{\rm r})}_{\rm res}\rangle$) simply identical with the vectors $|l_1,\dots,l_N\rangle$ ($|L_1,\dots,L_{N_{\rm r}}\rangle$). To each member (consisting of a copy of the system and of a copy of the reservoir) of the ensemble there corresponds therefore the Hilbert space $\mathcal{H}^{(N)}_{\rm sys}\otimes\mathcal{H}^{(N_{\rm r})}_{\rm res}$ which can be viewed as spaned either by the vectors

$$|l_1,\dots,l_N\rangle\otimes|L_1,\dots,L_{N_{\rm r}}\rangle\,,$$

or, equivalently, by the vectors $|n^{(N)}\rangle\otimes|n^{N_{\rm r}}_{\rm res}\rangle$. It is this second basis in terms of which the assignement of the statistical weight to the supersystems of the ensemble is given according to the principle of a priori equal probabilities by introducing the set of $N_{\rm tot}$ statistical operators $\hat{\rho}^{(N,N_r)}_{\rm Micro}$, each acting in one of the product Hilbert spaces $\mathcal{H}^{\rm sys}_{(N)}\otimes\mathcal{H}^{\rm res}_{(N_{\rm r})}$ and having the form

$$\hat{\rho}^{(N,N_r)}_{\rm Micro}={\rm Const.}\sum_{\substack{n^{(N)},\,n^{(N_{\rm r})}_{\rm res}\\ E_{\rm tot}\leq E^{\rm sys}_{n^{(N)}}+E^{\rm res}_{n^{(N_r)}_{\rm res}}\leq E_{\rm tot}+\Delta E}}\left(|n^{(N)}\rangle\otimes|n^{(N_{\rm r})}_{\rm res}\rangle\right)\left(\langle n^{(N_{\rm r})}_{\rm res}|\otimes\langle n^{(N)}|\right).$$

in which

$${\rm Const.}^{-1}=\sum_{N=0}^{N_{\rm tot}}\left(\begin{array}{c}\text{number of states } |n^{(N)}\rangle\otimes|n^{(N_{\rm r})}_{\rm res}\rangle\\ \text{with } E_{\rm tot}\leq E^{\rm sys}_{n^{(N)}}+E^{\rm res}_{n^{(N_r)}_{\rm res}}\leq E_{\rm tot}+\Delta E\end{array}\right). \tag{323}$$

The set of the statistical operators $\hat{\rho}^{(N)}$ (each acting in the N-particle Hilbert space $\mathcal{H}^{(N)}_{\rm sys}$ of the Grand Canonical Ensemble pertaining to the system alone is now obtained by taking traces of the operators $\hat{\rho}^{(N,N_r)}_{\rm Micro}$ over the Hilbert spaces $\mathcal{H}^{(N_{\rm r})}_{\rm res}$. Each such a trace gives the sum of the operators $|n^{(N)}\rangle\langle n^{(N)}|$ (acting in $\mathcal{H}^{(N)}_{\rm sys}$) weighted with the number of microstates of the reservoir corresponding to the number $N_r=N_{\rm tot}-N$ of particles in it and its energy in the interval $(E_{\rm tot}-E^{\rm sys}_{n^{(N)}},\,E_{\rm tot}+\Delta E-E^{\rm sys}_{n^{(N)}})$. The resulting operator can be, therefore, written in the form

$$\hat{\rho}^{(N)}\propto\sum_{n_{(N)}}|n_{(N)}\rangle\langle n_{(N)}|\exp\left\{\frac{1}{k_{\rm B}}\,S^{\rm res}(E_{\rm tot}-E^{\rm sys}_{n^{(N)}},\,N_{\rm tot}-N)\right\}.$$

Expanding now as usually the argument of the exponent

$$S^{\rm res}(E_{\rm tot}-E_{n_{(N)}},\,N_{\rm tot}-N)=S^{\rm res}(E_{\rm tot},N_{\rm tot})-\frac{1}{T}\,E^{\rm sys}_{n_{(N)}}+\frac{\mu}{T}\,N+\dots,$$

and discarding higher order terms of the expansion on account on the fact that they vanish in the limit of infinitely big reservoir ($N_{\rm tot}\to\infty$, $E_{\rm tot}\to\infty$, $E_{\rm tot}/N_{\rm tot}$ fixed), or modify the distribution insignificantly if the reservoir is very large but not really infinite, one obtains

$$\hat\rho^{(N)}\propto e^{\mu N/k_{\rm B}T}\sum_{n^{(N)}}e^{-E^{\rm sys}_{n^{(N)}}/k_{\rm B}T}\,|n^{(N)}\rangle\langle n^{(N)}|\equiv e^{\mu N/k_{\rm B}T}\,e^{-\hat{\mathscr{H}}^{(N)}_{\rm sys}/k_{\rm B}T}\,. \tag{324}$$

In this way the statistical operator $\hat\rho^{(N)}$ acting $\mathcal{H}^{(N)}_{\rm sys}$ is cast in the operator form, indepdendent of the particular choice of the basis in the N-particle Hilbert space. The constant of proportionality, the Grand Canonical statistical sum $\Xi_{\rm stat}$, is given by

$$\begin{aligned}\Xi_{\rm stat}&=\sum_{N=0}^{\infty}e^{\mu N/k_{\rm B}T}\sum_{n^{(N)}}e^{-E^{\rm sys}_{n^{(N)}}/k_{\rm B}T}\\&=\sum_{N=0}^{\infty}e^{\mu N/k_{\rm B}T}\,{\rm Tr}_{\mathcal{H}^{(N)}_{\rm sys}}\left(e^{-\hat{\mathscr{H}}^{(N)}_{\rm sys}/k_{\rm B}T}\right)\equiv\sum_{N=0}^{\infty}e^{\mu N/k_{\rm B}T}\,Z_{\rm stat}(T,V,N)\,.\end{aligned} \tag{325}$$

Again, the trace in the penultimate formula for $\Xi_{\rm stat}$ is independent of the choice of the basis in $\mathcal{H}^{(N)}_{\rm sys}$ and can, for instance, be computed using the basis $|l_1,\dots,l_N\rangle$ which are eigenvectors of the free part of the Hamiltonian $\hat{\mathscr{H}}^{(N)}_{\rm sys}$. As in the classical case, the sum $\Xi_{\rm stat}(T,V,\mu)$, being related to the thermodynamic potential (we assume the system is simple) $\Omega(T,V,\mu)$ by the formula (319), contains the complete thermodynamical information about the considered system. All other characteristics of the system, like mean values of different observables O, their fluctuations around the mean values, their correlations, etc. can be computed using the statistical operator (324).

Formulation in the "big" Hilbert space

It is practical to advance the mathematical formalism a little bit by introducing the "big" Hilbert space constructed as the direct sum

$$\mathcal{H}=\oplus_{N=0}^{\infty}\mathcal{H}^{(N)}_{\rm sys}\,, \tag{326}$$

of N-particle Hilbert spaces of the system. The Hilbert space $\mathcal{H}^{(0)}_{\rm sys}$ corresponding to zero particles is here an artificially constructed one-dimensional vector space spanned by a single vector $|{\rm void}\rangle$. General vectors $|\Psi\rangle$ of $\mathcal{H}$ have, of course components in all subspaces $\mathcal{H}^{(N)}_{\rm sys}$

$$|\Psi\rangle=c_0|{\rm void}\rangle+\sum_{N=1}^{\infty}c_N|\Psi_{(N)}\rangle\,,$$

where each vector $|\Psi_{(N)}\rangle$ is a linear combination of the basis vectors (307) and the scalar product in $\mathcal{H}$ is introduced naturally by declaring that $\langle\Psi_{(N)}|\Psi_{(N')}\rangle=0$ if $N\neq N'$. It is the big Hilbert space $\mathcal{H}$ in which the creation and annihilation operators can be introduced. With each single-particle state $|l\rangle$ (belonging to $\mathcal{H}^{(1)}_{\rm sys}$) a pair of the conjugated to one another operators $a_l^\dagger$ (the creation operator) and a_l (the annihilation operator) is associated and their action is defined by[194]

$$\begin{aligned} a_l^\dagger|l_1,\dots,l_N\rangle &= |l,l_1,\dots,l_N\rangle\,,\\ a_l|l_1,\dots,l_N\rangle &= \sum_{j=1}^{N}\zeta^{j-1}\delta_{l,l_j}|l_1,\dots,(\text{no } l_j,\dots,l_N\rangle\,, \end{aligned} \tag{327}$$

where $\zeta=1$, if the particles are bosons and -1, if they are fermions. These definitions must be supplemented by the rule $a_l|\text{void}\rangle=0$. Thus $a_l^\dagger$ maps $\mathcal{H}_{\text{sys}}^{(N)}$ into $\mathcal{H}_{\text{sys}}^{(N+1)}$ and a_l maps $\mathcal{H}_{\text{sys}}^{(N)}$ into $\mathcal{H}_{\text{sys}}^{(N-1)}$ (this is why the introduction of the big Hilbert space is necessary). It can be shown that the creation and annihilation operators satisfy the commutation rules

$$[a_{l'},\,a_l^\dagger]=\delta_{l'l}\,,\qquad [a_{l'},\,a_l]=[a_{l'}^\dagger,\,a_l^\dagger]=0\,, \tag{328}$$

if the particles are bosons and the anticommutation rules

$$\{a_{l'},\,a_l^\dagger\}=\delta_{l'l}\,,\qquad \{a_{l'},\,a_l\}=\{a_{l'}^\dagger,\,a_l^\dagger\}=0\,, \tag{329}$$

where $\{\cdot,\,\cdot\}$ is the anticommutator, if the particles are fermions. In the occupation number representation $|n_1,n_2,\dots\rangle$ of the basis vectors (307) the action of the bosonic operators takes the familiar (at least to some) form

$$\begin{aligned} a_l^\dagger|n_1,\dots,n_l,\dots\rangle &= \sqrt{n_l+1}\,|n_1,\dots,n_l+1,\dots\rangle\,,\\ a_l\,|n_1,\dots,n_l,\dots\rangle &= \sqrt{n_l}\,|n_1,\dots,n_l-1,\dots\rangle\,. \end{aligned} \tag{330}$$

The action of the fermionic operators in this representation is less familiar

$$\begin{aligned} a_l^\dagger|n_1,\dots,n_l,\dots\rangle &= \begin{cases} 0 & \text{if}\quad n_l=1\\ \eta|n_1,\dots,1,\dots\rangle & \text{if}\quad n_l=0 \end{cases}\\ a_l\,|n_1,\dots,n_l,\dots\rangle &= \begin{cases} \eta|n_1,\dots,0,\dots\rangle & \text{if}\quad n_l=1\\ 0 & \text{if}\quad n_l=0 \end{cases}\,. \end{aligned} \tag{331}$$

η is here some phase factor which we will not need (see my Lectures on Quantum Field Theory).

Usefulness of the introduced formalism stems from the fact that any operator acting in the individual N-particle Hilbert spaces $\mathcal{H}_{\text{sys}}^{(N)}$, like the free Hamiltonian of the form (309) as well as operators representing interactions (binary or even multibody) of particles can be expressed through the set of the creation and annihilation operators (associated with any basis $|l\rangle$ of the single-particle space $\mathcal{H}^{(1)}$) and at the same time and this promotes it to the operator acting in the entire big space $\mathcal{H}$ (326). In particular, introducing the particle number operator acting in the big Hilbert space $\mathcal{H}$ by the formula

$$\hat{\mathscr{N}}=\sum_{N=0}^{\infty}N\,\hat{1}^{(N)}=\sum_l a_l^\dagger a_l\,, \tag{332}$$

---

[194] Actually the action of a_l follows form the action of $a_l^\dagger$ and the requirement that a_l be the Hermitian conjugate of $a_l^\dagger$ in the sense of the definition (244).

where $\hat{1}^{(N)}$ are unit operators of the N-particle Hilbert spaces and expressing $\hat{\mathscr{H}}^{(N)}$ through the creation and annihilation operators it is possible to write the statistical operator of the Grand Canonical Ensemble in the concise form (we drop the subscript "sys")

$$\hat{\rho} = \frac{1}{\Xi_{\rm stat}}\exp\Big\{-\left(\hat{\mathscr{H}} - \mu\hat{\mathscr{N}}\right)/k_{\rm B}T\Big\}\,, \tag{333}$$

and the statistical sum $\Xi_{\rm stat}$ as the single trace

$$\Xi_{\rm stat} = {\rm Tr}_{\mathcal{H}}\Big(e^{-(\hat{\mathscr{H}}-\mu\hat{\mathscr{N}})/k_{\rm B}T}\Big)\,, \tag{334}$$

over the entire big space $\mathcal{H}$. In fact, there is a subtle point here related to the fact that the big Hilbert space $\mathcal{H}$ is nonseparable and different bases $|l\rangle$ of one-particle spaces $\mathcal{H}^{(1)}$ may lead to different results - the art then is to choose the right basis of $\mathcal{H}$ in which to perform actual computations corresponding to the real situation.

This can be illustrated most simply on the example of N distinguishable spins $s = \frac{1}{2}$ (forming e.g. a D-dimensional lattice). The natural basis of states of such a system is represented by the vectors (we assume spins are somehow ordered from 1 to N) $|\sigma_1,\dots,\sigma_N\rangle = |\sigma_1\rangle\otimes\dots\otimes|\sigma_N\rangle$ which are tensor products of the state-vectors $|\sigma_i\rangle = |\pm\rangle$ forming a basis of the single-spin space. In the space of a single spin one can however equally well take as the basis the vectors $|\sigma_i\rangle^\theta$: $|+\rangle^\theta = \cos(\theta/2)|+\rangle + \sin(\theta/2)|-\rangle$ and $|-\rangle^\theta = -\sin(\theta/2)|+\rangle + \cos(\theta/2)|-\rangle$ with an arbitrary angle θ corresponding to a different orientation of the spin quantization axis.[195] If the vectors $|\sigma_1,\dots,\sigma_N\rangle^\theta = |\sigma_1\rangle^\theta\otimes\dots\otimes|\sigma_N\rangle^\theta$ are taken as the basis of the N-spin Hilbert space, in the limit $N=\infty$ all vectors of the basis $|\sigma_1,\dots\sigma_N\rangle^\theta$ will have zero scalar products with all vectors of the basis $|\sigma_1,\dots,\sigma_N\rangle$. For instance

$${}^\theta\langle+,\dots,+|+,\dots,+\rangle = \lim_{N\to\infty}(\cos\theta/2)^N = 0\,,$$

etc., if $\theta\neq 0$. Therefore in the limit $N=\infty$ no one vector of the first basis can be expressed as a linear combination of vectors of the second one (and the other way around).

To explain the implications of this, consider a system of N spins interacting with one another which at low temperatures exhibits (because of the interactions of spins) spontaneous (that is, in the absence of any extrnal magnetic field) magnetization, i.e. a nonzero mean value of the total spin S, which formally should be obtained as[196]

$$\overline{S^a} = {\rm Tr}\Big(\hat{\rho}\,\hat{S}^a\Big)\,,\quad \hat{S}^a = \sum_{i=1}^{N}\frac{\hbar}{2}\,\sigma^a_i\,,\quad a = x,y,z\,, \tag{335}$$

---

[195] Of course, one could also take different quantization axes for different spins.

[196] We chose to work within the Canonical Ensemble here. The problem we consider here is not specific to the Grand Canonical Ensemble. It arises whenever the number of dynamical degrees of freedom becomes infinite.

with $\hat{\rho} = Z_{\rm stat}^{-1}\exp(-\hat{\mathscr{H}}/k_{\rm B}T)$, where $\hat{\mathscr{H}}$ is the invariant (by assumption) with respect to rotations Hamiltonian of N spins and $Z_{\rm stat} = {\rm Tr}\exp(-\hat{\mathscr{H}}/k_{\rm B}T)$. However, as long as N is finite, the values of $\overline{S^a}/N$ given by formula (335) are zero because of the assumed rotational invariance. Of course, in Nature N is also finite, but the residual interactions with the surrounding, not accounted (by the very definition of the ensemble!) in the Hamiltonian $\hat{\mathscr{H}}$ used for statististical physics computations, always break the rotational invariance of the real system. Thermal fluctuations, necessarily present, do tend to erase the mean value of the total spin but because the system is large (macroscopic), the probability that they move simultaneously (almost) all spins - so that the change of the total spin of the system is done at low energy cost (gradual fliping of different spins would cost a large amount of energy which the system - if not adiathermally isolated in the macroscopic sense - would have to absorb from the surrounding) - from a configuration determined by a first random external perturbation (or rather by the way the real system has been prepared) grows with N, the flip of the total spin is extremely improbable and never happens during the measurement of the magnetization. This is the reason why the measured magnetization of the real system is nonzero (if the temperature is suffeciently low). To reflect this state of affairs in the statistical approach, one takes the thermodynamic limit selecting first a particular basis $|\sigma_1, \dots \sigma_N\rangle^\theta$; in the limit $N = \infty$ this has the effect that the formula (335) gives a nonzero magnetization precisely owing to the orthogonality of all states $|\sigma_1, \dots \sigma_N\rangle^{\theta'}$ with $\theta' \neq \theta$. This means that the trace in this formula effectively gets restricted to only one of infinitely many orthogonal Fock spaces which in the limit $N = \infty$ are spanned by the bases $|\sigma_1, \dots \sigma_N\rangle^\theta$ with different angles θ. The direction of the mean magnetization is in this way selected “by hands”, just by taking the thermodynamic limit using one particular out of many possible bases.[197] This can be improved by placing the system in a constant magnetic field which singles out a direction in space and therefore makes the energy cost of configurations with the total spin not aligned with the magnetic field very high, infinite in the limit $N = \infty$, and this suppresses their contributions to the trace in (335). In this way the applied field automatically selects one particular base $|\sigma_1, \dots \sigma_N\rangle^\theta$. In this approach the magnetization is nonzero even at finite N because of the external field breaking explicitly the rotational invariance and he limit of zero external magnetic field is to be taken after the thermodynamic limit $N \to \infty$ (the thermodynamic and the $N \to \infty$ limits do not commute).

Nointeracting particles. (Mythical quantum statistics)

To get acquainted with the formalism let us first consider a system consisting of mutually noninteracting particles (molecules). The Hamiltonian of the N-particle system of the form (309) and the particle number operator (332) when written in terms of the creaction and annihilation operators a_l and $a_l^\dagger$ associated with the eigenvectors $|l\rangle$ of the single

---

[197] It is interesting to remark that if the system is treated within the Microcanonical Ensemble (the real system is treated as macroscopically isolated) obtaining a nonzero value of the total spin statistically requires breaking the “ergodicity” that is, only microstates leading to one direction of the total spin are ascribed a nonzero probability; microstates corresponding to opposite total spin orientation must be excluded “by hands”.

particle Hamiltonian $\hat{\mathscr{H}}^{(1)}$ (this Hamiltonian may include also single-particle interaction with e.g. an external electric or magnetic field or of an external biding potential) take the forms (which, as said, promote them to operators acting in the entire big space $\mathcal{H}$)

$$\hat{\mathscr{H}} = \sum_l \varepsilon_l \, a_l^\dagger a_l \,, \qquad \hat{\mathscr{N}} = \sum_l a_l^\dagger a_l \,. \tag{336}$$

The grand statistical sum can in this case be easily computed using in the formula (334) the occupation number representation of the common eigenvectors (which form, at least formally - this is a somewhat subtle issue - the basis of $\mathcal{H}$) of the operators (336)

$$\Xi_{\rm stat} = \sum_{N=0}^{\infty} \sum_{n_1=0}^{n^{\rm max}} \sum_{n_2=0}^{n^{\rm max}} \dots \delta_{N,\sum_l n_l} \langle n_1, n_2, \dots | \exp\left(-\beta \sum_l (\varepsilon_l - \mu) a_l^\dagger a_l\right) | n_1, n_2, \dots \rangle \,,$$

where $n^{\rm max} = \infty$, if the considered particles (molecules) are bosons and $n^{\rm max} = 1$, if they are fermions. The sum over N effectively removes the troublesome Kronecker deltas and

$$\Xi_{\rm stat} = \sum_{n_1=0}^{n^{\rm max}} \sum_{n_2=0}^{n^{\rm max}} \dots \langle n_1, n_2, \dots | \exp\left(-\beta \sum_l (\varepsilon_l - \mu) a_l^\dagger a_l\right) | n_1, n_2, \dots \rangle \,.$$

Therefore,

$$\Xi_{\rm stat} = \prod_l \left(\sum_{n=0}^{n^{\rm max}} e^{-\beta(\varepsilon_l - \mu)n} \right) = \left\{ \begin{array}{ll} \prod_l [1 - \exp(-\beta(\varepsilon_l - \mu)]^{-1}, & \text{bosons} \\ \prod_l [1 + \exp(-\beta(\varepsilon_l - \mu)], & \text{fermions} \end{array} \right. . \tag{337}$$

Owing to the Pauli exclusion principle (built in into the formalism), the sum in the bracket reduces, if the particles are fermions, to two terms only. The analogous sum in the case of free bosons requires instead performing the summation of the infinite geometric series. Its convergence therefore imposes on the chemical potential μ the condition

$$\mu < \min_l(\varepsilon_l) \qquad \text{bosons.} \tag{338}$$

Using the statistical operator one can also compute the mean occupation $\overline{n_l}$ of a single-particle state $|l\rangle$. The observable n_l is represented by the operator $a_l^\dagger a_l$ and

$$\overline{n_l} = \frac{1}{\Xi_{\rm stat}} \, {\rm Tr}\Big(e^{-\beta(\hat{\mathscr{H}} - \mu \hat{\mathscr{N}})} \, a_l^\dagger a_l \Big) \,.$$

In the occupation number representation this reduces to

$$\overline{n_l} = \frac{1}{\Xi_{\rm stat}} \left(\sum_{n_l=0}^{n^{\rm max}} n_l \, e^{-\beta(\varepsilon_l - \mu)n_l} \right) \prod_{l' \neq l} \left(\sum_{n_{l'}=0}^{n^{\rm max}} e^{-\beta(\varepsilon_l - \mu)n_{l'}} \right) .$$

All the sums of this product, except the first one, cancel between the numerator and the denominator and the result are the celebrated Bose-Einstein and Fermi-Dirac distributions:[198]

$$\begin{aligned}\overline{n_l} &= \left[e^{(\varepsilon_l-\mu)/k_{\rm B}T} - 1\right]^{-1} \qquad \text{bosons,}\\ \overline{n_l} &= \left[e^{(\varepsilon_l-\mu)/k_{\rm B}T} + 1\right]^{-1} \qquad \text{fermions.}\end{aligned} \tag{339}$$

Many quantities characterizing equilibrium states **of noninteracting systems** can be expressed directly in terms of these distributions. For instance, the potential $\Omega = -k_{\rm B}T\ln\Xi_{\rm stat}$ can be written (c.f. (337)) as

$$\begin{aligned}\Omega &= -k_{\rm B}T\sum_l \ln(1+\overline{n_l}) \qquad \text{bosons,}\\ \Omega &= k_{\rm B}T\sum_l \ln(1-\overline{n_l}) \qquad \text{fermions,}\end{aligned} \tag{340}$$

and the mean energy as

$$U = \overline{E} = {\rm Tr}_{\mathcal{H}}\left(\hat{\rho}\hat{\mathscr{H}}\right) = \sum_l \overline{n_l}\,\varepsilon_l\,, \tag{341}$$

(bosons and fermions alike). We leave the calculations of entropy of the system ($S = -(\partial\Omega/\partial T)_{V,\mu}$) in terms of the mean occupation numbers as well as the calculation of the mean quadratic fluctuation of the occupation number n_l as good exercises.

Gas of noninteracting particles in the quantum regime.
We will now apply the Grand Canonical Ensemble to the system of mutually noninteracting particles (bosons or fermions).

As usually we assume that N particles occupy the volume $V = L^3$ and that the single-electron Hilbert space is spanned by the vectors $|l\rangle = |\mathbf{p}, \sigma\rangle$ which are eigenvectors of the single particle Hamiltonian $\hat{\mathscr{H}}^{(1)}$ and correspond to its eigenvalues (energies of individual particles) $\varepsilon_{\mathbf{p}} = \mathbf{p}^2/2m$ (independent of the spin label σ), where $\mathbf{p} = (2\pi\hbar/L)\,\mathbf{n}$, with $\mathbf{n}$ being three-vectors of integer components. The label σ runs over $g_s = 2s+1$ values $-s, \dots, +s$, where s is half-integer if the particles are fermions ($s = \frac{1}{2}$ and $g_s = 2$ if they are electrons) and integer, if they are bosons. The statistical sum is given by the formulae (337) in which now the products over l stand for the products over $\mathbf{p}$ and σ. After the standard replacement

$$\sum_{\mathbf{p},\sigma} \to \frac{V}{(2\pi\hbar)^3}\sum_\sigma \int d^3\mathbf{p}\,,$$

---

[198] The Bose-Einstein distribution has been first derived by a polish theoretical physicist Władysław Natanson (1864 - 1937) in 1911 with the help of the combinatoric reasoning and by appealing to the indistinguishability of particles (L. Natanson, Physikalishe Zeitschrift **12** 659, (1911)). I thank prof. K. Byczuk and prof. P. Sułkowski for recalling this fact.

the thermodynamical potential $\Omega(T,V,\mu)=U-TS-\mu N=-k_{\rm B}T\ln\Xi_{\rm stat}(T,V,\mu)$ takes the form

$$\Omega(T,V,\mu)=\pm k_{\rm B}T\,g_sV\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\,\ln\!\big(1\mp z\,e^{-\beta\varepsilon_{\mathbf{p}}}\big)\,,$$

in which $z\equiv\exp(\mu/k_{\rm B}T)=\exp(\beta\mu)$, is the activity and the upper and lower signs correspond to bosons and fermions respectively. The potential $\Omega(T,V,\mu)$, which is given in this way as a function of its natural variables, encodes in it the complete thermodynamical (i.e. not taking into account statististical fluctuations) information about the system.

Differentiating the (minus) potential Ω with respect to μ at fixed T and V one obtains the mean number N of particles in the system (since this is identified with the thermodynamic quantity, in what follows we will suppress the bar over it). The same quantity can be however obtained as the sum over l (replaced by the integral) of the mean occupation numbers $\overline{n_l}$ given by the celebrated formulae (339). Yet one more quantity of interest which can be obtained by appropriately combining derivatives of the potential $\Omega(T,V,\mu)$ but is much easier obtained as the mean ${\rm Tr}(\hat\rho\,\hat{\mathscr{H}})$ over the ensemble which directly leads to the expression (341) is the system's internal energy U. These three - counting also the formula for $\Omega(V,T,\mu)=-V\,p(T,\mu)$ itself - results are most conveniently represented in the forms

$$\begin{aligned}\frac{p}{k_{\rm B}T}&=\mp g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\,\ln\!\big(1\mp z\,e^{-\beta\varepsilon_{\mathbf{p}}}\big)\,,\\ \frac{N}{V}\equiv\frac{1}{v}&=g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\,\frac{1}{z^{-1}e^{\beta\varepsilon_{\mathbf{p}}}\mp1}\,,\\ \frac{U}{V}&=g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\,\frac{\varepsilon_{\mathbf{p}}}{z^{-1}e^{\beta\varepsilon_{\mathbf{p}}}\mp1}\,,\end{aligned}\tag{342}$$

(again, upper signs refer to bosons, the lower ones - to fermions). In order to express thermodynamical quantities in the conventional way as functions of the variables T, V i N, one has to solve the middle equation with respect to μ, that is to determine from this equation the chemical potential as a function of T and $V/N=v$. This is of course not feasible analytically[199] and one has to resort to different approximations (expansions) valid in two different regimes: one in the limit $z\ll1$, i.e. large negative chemical potential μ (here one is interested in the corrections to the Boltzmann approximation discussed in Lecture XII brought in by the exact accounting for the symmetry/antisymmetry of the system's state-vectors) and another one in the opposite regime $k_{\rm B}T\to0$ (in which quantum effects should lead to vanishing of the system's entropy and heat capacity).

---

[199] It is in two dimensions (in the nonrelativistic case, i.e. when $\varepsilon_{\mathbf{p}}=\mathbf{p}^2/2m$), owing to the identity

$$\frac{1}{z^{-1}e^x\mp1}=\pm\frac{d}{dx}\ln\!\big(1\mp z\,e^{-x}\big)\,.$$

One then obtains $\mu=k_{\rm B}T\ln[-1+\exp(2\pi\hbar^2n/g_smk_{\rm B}T)]$, if particles are fermions and $\mu=k_{\rm B}T\ln[1-\exp(-2\pi\hbar^2n/g_smk_{\rm B}T)]$, if they are bosons; in both these formulae n is the surface density of particles.

Some information about the system's thermodynamics contained in the formulae (342) (in fact in the statistical sum $\Xi_{\rm stat}$) can however be extracted rigorously and we discuss these first.

To this end, going over to integrating over the energy with the help of the rule (that is performing the integral over the directions of the momentum $\mathbf{p}$)

$$d^3\mathbf{p} \longrightarrow 4\pi\sqrt{2}\,m^{3/2}\,d\varepsilon\,\sqrt{\varepsilon}\,, \tag{343}$$

(valid as long as we treat particles as nonrelativistic), we represent the first one of the three formulae (342) in the form

$$\frac{p}{k_{\rm B}T} = \mp g_s\,\frac{\sqrt{2}\,m^{3/2}}{(2\pi\hbar)^3}\,4\pi\int_0^\infty d\varepsilon\,\sqrt{\varepsilon}\ln\!\left(1\mp z\,e^{-\beta\varepsilon}\right),$$

and integrate by parts:

$$\frac{p}{k_{\rm B}T} = \mp g_s\,\frac{\sqrt{2}\,m^{3/2}}{(2\pi\hbar)^3}\,4\pi\left\{\frac{2}{3}\,\varepsilon^{3/2}\ln\!\left(1\mp z\,e^{-\beta\varepsilon}\right)\Big|_0^\infty \mp\frac{2}{3}\int_0^\infty d\varepsilon\,\varepsilon^{3/2}\,\frac{\beta}{z^{-1}\,e^{\beta\varepsilon}\mp 1}\right\}.$$

The boundary term vanishes: in the lower limit because of the factor $\varepsilon^{3/2}$ and in the upper one, because the expression under the logarithm approaches unity (exponentially fast). After taking into account that $\beta = 1/k_{\rm B}T$ (the factors $1/k_{\rm B}T$ on both sides of the equality can therefore by removed) one obtains for the pressure the expression in which the integral on the right hand side is identical with the integral obtained on the right hand side of the third of the formulae (342) after applying to it the rule (343). Thus we rigorously (without making any approximations or expansions) arrive at the conclusion that the quantum gas of noninteracting nonrelativistic particles, whether they are fermions or bosons,[200] satisfies the relation

$$p\,V = \frac{2}{3}U\,, \tag{344}$$

identical with the one satisfied by the classical perfect gas that is, a gas of mutually noninteracting structureless particles treated either according the the rules of classical statistical mechanics or according to the quantum statistical mechanics but in the Boltzmann approximation which ignores effects of the (anti)symmetrization of the state vectors.

Another result pertaining to the quantum gas of noninteracting nonrelativistic particles (bosons or fermions) which can be obtained rigorously is the equation of its adiabatic (reversible adiathermal) change (the isentrop equation). Going over in the integral

$$\Omega = -p\,V = -\frac{2}{3}\,g_s\,\frac{\sqrt{2}\,m^{3/2}}{(2\pi\hbar)^3}\,4\pi\,V\int_0^\infty d\varepsilon\,\frac{\varepsilon^{3/2}}{z^{-1}\,e^{\beta\varepsilon}\mp 1}\,, \tag{345}$$

---

[200] Actually, in the case of bosons the derivation given here is not complete for very low temperatures because it neglects the possibility of the Bose-Einstein condensation (Lecture XIV). Still, the final result (344) turns out to hold also in the presence of the condensation.

to integration over the dimensionless variable $x = \beta\varepsilon \equiv \varepsilon/k_{\rm B}T$, the potential Ω can be written in the general form

$$\Omega(T, V, \mu) = V\,(k_{\rm B}T)^{5/2} f(\mu/k_{\rm B}T)\,, \tag{346}$$

in which $f(\cdot)$ is some function of a real argument. This form makes it clear that $\Omega(T, V, \mu)$ is a homogeneous function of order $5/2$ of its arguments T and μ that is, that it satisfies the relation

$$\Omega(\lambda T,\, V,\, \lambda\mu) = \lambda^{5/2}\,\Omega(T,\, V,\, \mu)\,.$$

From the general properties of homogeneous functions it then follows that the derivatives of $\Omega(T, V, \mu)$ with respect to T and μ, that is the entropy S and the (mean) number of particles N, must be homogeneous functions of order $3/2$ (of the same variables), and can, therefore, be written in general in the forms

$$S(T, V, \mu) = V\,(k_{\rm B}T)^{3/2}\, g(\mu/k_{\rm B}T)\,,$$
$$N(T, V, \mu) = V\,(k_{\rm B}T)^{3/2}\, h(\mu/k_{\rm B}T)\,,$$

with $h(\cdot)$ and $g(\cdot)$ some functions of a real argument. Factorization of the volume V follows in all these three cases from the formula (345) and expresses the extensiveness of Ω, S and N. From these formulae it readily follows that in an adiabatic (reversible, adiathermal) change in which S and N do not change, separately constant must be the combinations $\mu/k_{\rm B}T$ and $V\,(k_{\rm B}T)^{3/2}$. Indeed, the entropy per particle, $s = S/N$, is a function of the single variable $\mu/k_{\rm B}T$ only and since s stays constant, constant must also be $\mu/k_{\rm B}T$. The constancy of $V\,(k_{\rm B}T)^{3/2}$ then follows from constancy of S and N. From (346) written in the form

$$p\,V = -V\,(k_{\rm B}T)^{5/2} f(\mu/k_{\rm B}T) = \text{Const.}\, V\,(k_{\rm B}T)^{5/2}\,,$$

it then follows that in such a change $pV = \text{Const.}' k_{\rm B}T$ and this combined with the constancy of $V\,(k_{\rm B}T)^{3/2}$ implies that $p\,V^{5/3} = \text{Const.}$, just as in the case of an adiabatic change of a classical gas of (structureless) particles. It should be stressed that the adiabat equation of the quantum perfect gas of nonrelativistic particles is identical with the one of the classical perfect gas despite that[201] it is neither true that its heat capacity is independent of the temperature T nor that $C_p = C_V + Nk_{\rm B}$.

Corrections to the perfect gas equation of state

We now consider the regime in which $z \ll 1$ (i.e. large negative chemical potential μ). This is the limit in which the Boltzmann approximation works and leads to the perfect gas equation of state[202] $p\,V = Nk_{\rm B}T$. We will compute the correction to this equation

---

[201] For those who still do not remember it, let us recall that the form $pV^{\kappa} = \text{const.}$ of the adiabat equation of the classical perfect gas relies on the assumption of constancy of the its heat capacity C_V; the exponent $\kappa = 5/3$ is then just the ratio C_p/C_V, when $C_V = \frac{3}{2}Nk_{\rm B}$ (by the Meyer's relation $C_p = C_V + Nk_{\rm B}$).

[202] Let us recall, that within classical statistical mechanics as well as within the Boltzmann approximation to the quantum canonical statistical sum, this is the equation of state of any gas of mutually noninteracting molecules, not only of the gas of structureless particles of molar heat capacity independent of the temperature.

of state resulting from the proper treatement of the requirement of (anti)symmetrization of the system's state vectors. For definiteness we will consider the gas of fermions but the analogous steps can be applied to the equations pertaining to the gas of bosons (the necessary auxiliary functions will be defined when we consider the phenomenon of Bose-Einstein condensation) with the final leading correction to the equation of state differing only by the sign.

To determine the equation of state we should solve the second one of the equations (342) to obtain the chemical potential μ as a function of T and v and insert it into the first equation. As already said, this can be done only by using appropriate expansions. The one applicable in the limit $|z| \ll 1$ is constructed as follows. After performing in the formulae (342) the integrations over the directions of the momentum $\mathbf{p}$ one writes the expressions remaining on the right hand sides as the integrals over the variable $\xi = |\mathbf{p}|/\sqrt{2mk_{\rm B}T}$. The two equations then read:

$$\begin{aligned}\frac{p}{k_{\rm B}T} &= g_s\,\frac{4\pi}{(2\pi\hbar)^3}\,(2mk_{\rm B}T)^{3/2}\int_0^\infty d\xi\,\xi^2\ln\!\Big(1+z\,e^{-\xi^2}\Big)\,,\\ \frac{1}{v} &= g_s\,\frac{4\pi}{(2\pi\hbar)^3}\,(2mk_{\rm B}T)^{3/2}\int_0^\infty d\xi\,\xi^2\frac{1}{1+z^{-1}e^{\xi^2}}\,.\end{aligned}$$

"Borrowing" the factors $\sqrt{\pi}$ one can rewrite these formulae as

$$\frac{p}{k_{\rm B}T} = g_s\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} f_{5/2}(z)\,, \tag{347}$$

$$\frac{1}{v} = g_s\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} f_{3/2}(z)\,, \tag{348}$$

thereby definig rigorously (using the language of quantum field theory one would say "nonperturbatively") two functions (f, because of fermions; analogous functions used to work out the properties of bosons will be denoted $b_p(z)$) $f_{5/2}(z)$ and $f_{3/2}(z)$;

$$\begin{aligned}f_{5/2}(z) &= \frac{4}{\sqrt{\pi}}\int_0^\infty d\xi\,\xi^2\ln\!\Big(1+z\,e^{-\xi^2}\Big) = \frac{2}{\sqrt{\pi}}\int_0^\infty dx\,\sqrt{x}\ln\!\big(1+z\,e^{-x}\big)\,,\\ f_{3/2}(z) &= \frac{4}{\sqrt{\pi}}\int_0^\infty d\xi\,\xi^2\,\frac{1}{1+z^{-1}e^{\xi^2}} = \frac{2}{\sqrt{\pi}}\int_0^\infty dx\,\frac{\sqrt{x}}{1+z^{-1}e^{x}}\,.\end{aligned}$$

An alternative form of $f_{5/2}(z)$ is[203]

$$f_{5/2}(z) = \frac{2}{3}\frac{4}{\sqrt{\pi}}\int_0^\infty d\xi\,\xi^4\frac{1}{1+z^{-1}e^{\xi^2}}\,.$$

In general the family of functions $f_p(z)$ is defined by the integrals

$$f_p(z) = \frac{1}{\Gamma(p)}\int_0^\infty dx\,x^{p-1}\,\frac{1}{1+z^{-1}e^{x}} = \frac{1}{\Gamma(p)}\int_0^\infty dx\,x^{p-1}\,\frac{z\,e^{-x}}{1+z\,e^{-x}}\,. \tag{349}$$

---

[203] The first form of $f_{5/2}(z)$ (the one with the logaritm) can be reduced to this one by integrating by parts in the same way as in showing that $p\,V = (2/3)U$.

It is straightforward to write their series expansions in powers of z:

$$f_p(z) = \frac{1}{\Gamma(p)}\int_0^\infty dx\, x^{p-1}\sum_{l=1}^\infty (-1)^{l+1} z^l\, e^{-lx} = \sum_{l=1}^\infty (-1)^{l+1} z^l\, \frac{1}{\Gamma(p)}\int_0^\infty dx\, x^{p-1}\, e^{-lx}\,.$$

Rescaling now the integration variable, i.e. setting $lx = y$, one arrives, using the integral representation of $\Gamma(p)$, at the result[204]

$$f_p(z) = \sum_{l=1}^\infty \frac{(-1)^{l+1}}{l^p}\, z^l\,. \tag{350}$$

The series converges only for $|z| < 1$ (for $|z| \leq 1$ when $p > 1$), but the "nonperturbative" definition (349) of $f_p(z)$ allows to analytically continue the series beyond its domain of convergence.

Using the series (350) the ratio of the expressions: (347) for $p/k_{\rm B}T$ and (348) for $1/v$ can be easily expanded in powers of z:

$$\frac{p\, v}{k_{\rm B}T} = \frac{f_{5/2}(z)}{f_{3/2}(z)} = 1 + \frac{1}{2^{5/2}}\, z + \dots \tag{351}$$

To find the leading correction to the equation of state it suffices now to invert to the first order the relation

$$\frac{1}{v} = g_s\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} f_{3/2}(z) = g_s\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} z + \mathcal{O}(z^2)\,,$$

to determine z as a function of the variables T and v. This immediately leads to the result (λ_T is the thermal wavelength)

$$p\, v = k_{\rm B}T\left(1 + \frac{1}{2^{5/2}g_s}\left(\frac{2\pi\hbar^2}{mk_{\rm B}T}\right)^{3/2}\frac{1}{v} + \dots\right) = k_{\rm B}T\left(1 + \frac{1}{2^{5/2}g_s}\,\frac{\lambda_T^3}{v} + \dots\right). \tag{352}$$

The pressure of the quantum perfect gas of fermions is slightly greater than that of the classical perfect gas (at the same temperature and density $1/v$): fermions do not "like" to be close to one another which results, even in the absence of a genuine interaction, in an effective repelling which increses the pressure. It is therefore an effect of the Fermi-Dirac "statistics", that is of the underlying antisymmetry of state vectors representing the system of identical fermions.

---

[204] The first form of $f_{5/2}(z)$ leads, of course, to the same expansion: after expanding the logarithm

$$f_{5/2}(z) = \frac{4}{\sqrt{\pi}}\sum_{l=1}^\infty \frac{(-1)^{l+1}}{l}\, z^l \int_0^\infty d\xi\, \xi^2\, e^{-l\xi^2}\,.$$

the subsitution $\eta = \xi\sqrt{l}$ leads to an integral which is independent of l and precisely cancels the factor $4/\sqrt{\pi}$.

The derivation of the leading correction to the equation of state of the perfect gas of bosons proceeds analogously (one has only to replace the $f_p(z)$ functions by the $b_p(z)$ ones) and gives the same correction[205] as in (352) but with the opposite sign: bosons do "like" to "gather" (especially if illegally) which results in their effective attraction (even in the absence of a genuine interaction) which slightly decreases the pressure. This too is the effect of the "statistics", in this case of the Bose-Einstein one (of the symmetry of state vectors representing systems of identical bosons).

Relativistic theories

A characteriscic feature of interactions of relativistic particles is that their number is not conserved. Therefore the Grand Canonical Ensemble in the form presented above cannot be applied to systems of relativistic particles. Usually however models of quantum field theory - which are the proper framework in which to consider relativistic interactions - predict strict conservation (in addition to energy, momentum and angular momentum) of certain quantities, like e.g. the electric charge in the case of Quantum Electrodynamics or isospin, strangeness, etc. in the case of Quantum Chromodynamics - the theory of strong interactions, etc. These quantities, let denote them Q_b, generally called charges, are related to symmetries of the given model. The operators representing them in the theory's Hilbert space are Hermitian and form a closed algebra playing the roles of symmetry generators (much in the same way as the angular momentum operators generate rotations). Usually the symmetry generators do not commute, $[\hat{Q}_b, \hat{Q}_a] \neq 0$, but it is always possible to select a subset of them, $\hat{Q}_b$, $b = 1, \dots, r$, which forms the so-called Cartan subalgebra of mutually commuting generators. It is with the quantities represented by these operators that one associates the set of chemical potentials μ_b and constructs the Grand Canonical Ensemble representing the system in thermal contact with the heat bath with which it exchanges also the quantities Q_b, $b = 1, \dots, r$.

When reactions between relativistic particles are considered one has to appeal to the Callenian concept of virtual equilibrium states in which reactions are "frozen". It is then possible to ascribe chemical potentials to individual types of particles. When the reactions are next allowed, the final state is such that for every reaction $i_1 + i_2 + \dots \leftrightarrow i'_1 + i'_2 + \dots$ which can occur between relativistic particles in the system's way to equilibrium the relation

$$\mu_{i_1} + \mu_{i_2} + \dots = \mu_{i'_1} + \mu_{i'_2} + \dots \tag{353}$$

(obtained within the thermodynamical theory of chemical reactions) must be satisfied where μ_{i_1}, ... are the chemical potentials of the particles entering the reaction and $\mu_{i'_1}$, ... are the chemical potentials of the particles which are products of the reaction. The chemical potentials μ_i of particles of the given theory are not independent but must be fixed by the chemical potentials μ_b of the absolutely conserved quantities Q_b. They are given by the relations

$$\mu_i = \sum_{b=1}^{r} q_{ib}\, \mu_b\,, \tag{354}$$

---

[205] Of course, the factor g_s must be different in this case.

which ensure, since in every reaction the sum of the charges q_{ib} of the particles defined by $\hat{Q}_b|\mathbf{p}_i,\sigma_i\rangle = q_{ib}|\mathbf{p}_i,\sigma_i\rangle$, must be conserved (because Q_a are absolutely conserved quantities!), that all the relations (353) which pertain to the reactions possible (even if predicted to occur with a very low probability!) within a given quantum field theory model, are automatically satisfied (in a given quantum field theory model usually infinitely many reactions are possible).

In applications of statistical mechanics to relativistic theories usually the values of the densities $n_b = Q_b/V$ of the conserved quantities are fixed (the Grand Canonical Ensemble approach is only a technical mean to cope with their conservation). In terms of the densities n_i of the particles of the sort i they are given by

$$n_b = \sum_i n_i\, q_{ib}\,. \tag{355}$$

These are then the relations which in principle allow to determine the r chemical potentials μ_b, and, through (354), also the dependent chemical potentials of stable particles.

Systems of interacting particles (elements)

Nointeracting systems are rather artificial constructs, applicable at first sight to only very special physical situations; the most interesting physical phenomena are usually due to interactions correlating the behaviour of individual elements out of which real physical systems are made. Nevertheless, the importance and ubiquity of the results obtained without interactions, in particular of the Bose-Einstein and Fermi-Dirac distributions (i.e. of the "quantum statistics") stems from the fact that they constitute the first terms of the perturbative expansions by which realistic, interacting systems can be handled. It is here that the formalism of the Fock space and of the creation and annihilation operators coupled with the general quantum field theory methods fully shows its advantages. Here we only outline the lines along which systematic perturbative computations of properties of interacting statistical systems can be organized. We limit ourselves to the computation of the statistical sum $\Xi_{\rm stat}$ which contains the complete thermodynamical information about the considered system although computations of mean values of observables and of various correlation functions can be formulated similarly.

Suppose the Hamiltonian of the system of interacting elements consists of two terms $\hat{\mathscr{H}} = \hat{\mathscr{H}}_0 + \hat{\mathscr{V}}_{\rm int}$ with $\hat{\mathscr{H}}_0$ of the form (336) and $\hat{\mathscr{V}}_{\rm int}$ represents interactions coupling elements of the system. Such an interaction term can always be expressed through the creation and annihilation operators associated with the basis of the single-particle Hamiltonian $\hat{\mathscr{H}}_0^{(1)}$. For example, if the system consists of particles (or molecules the internal structure of which can be neglected) of mass m and spin s (bosons or fermions) enclosed in the box of volume $V = L^3$ and interacting with one another through the spin-independent potentials $\mathscr{V}(\mathbf{r}_i - \mathbf{r}_j)$, the Hamiltonian takes the form

$$\hat{\mathscr{H}} = \sum_{\mathbf{p},\sigma} \frac{\mathbf{p}^2}{2m}\, a^\dagger_{\mathbf{p},\sigma} a_{\mathbf{p},\sigma} + \frac{1}{2V}\sum_{\mathbf{q}} \tilde{\mathscr{V}}(\mathbf{q}) \sum_{\mathbf{p},\sigma}\sum_{\mathbf{p}',\sigma'} a^\dagger_{\mathbf{p}+\mathbf{q},\sigma} a^\dagger_{\mathbf{p}'-\mathbf{q},\sigma'} a_{\mathbf{p}',\sigma'} a_{\mathbf{p},\sigma}\,,$$

where $\tilde{\mathscr{V}}(\mathbf{q})$ is the Fourier transform of the potential $\mathscr{V}(\mathbf{r})$.

The perturbative expansion of the statistical sum $\Xi_{\rm stat} = {\rm Tr}_{\mathcal{H}}\{\exp(-\beta(\hat{\mathscr{H}}-\mu\hat{\mathscr{N}})\}$ can be formulated by defining the "interaction picture imaginary time" evolution operator[206]

$$\hat{\mathscr{U}}_I(\beta,0) = e^{\beta\hat{\mathscr{K}}_0}\,e^{-\beta\hat{\mathscr{K}}}\,, \tag{356}$$

in which $\hat{\mathscr{K}}_0 = \hat{\mathscr{H}}_0 - \mu\hat{\mathscr{N}}$ and $\hat{\mathscr{K}} = \hat{\mathscr{H}} - \mu\hat{\mathscr{N}}$. The operator $\hat{\mathscr{U}}_I(\beta,0)$ satisfies the differential equation

$$\frac{d}{d\beta}\,\hat{\mathscr{U}}_I(\beta,0) = -\hat{\mathscr{V}}^I_{\rm int}(\beta)\,\hat{\mathscr{U}}_I(\beta,0)\,, \tag{357}$$

in which

$$\hat{\mathscr{V}}^I_{\rm int}(\beta) = e^{\beta\hat{\mathscr{K}}_0}\,\hat{\mathscr{V}}_{\rm int}\,e^{-\beta\hat{\mathscr{K}}_0}\,,$$

with the "initial" condition $\hat{\mathscr{U}}_I(0,0) = \hat{1}$. The differential equation can be converted into the integral equation

$$\hat{\mathscr{U}}_I(\beta,0) = \hat{1} - \int_0^\beta d\beta'\,\hat{\mathscr{V}}^I_{\rm int}(\beta')\,\hat{\mathscr{U}}_I(\beta',0)\,,$$

into which the initial condition is automaically built in. One can now solve this integral equation iteratively (Banach's principle!) obtaining the expansion

$$\hat{\mathscr{U}}_I(\beta,0) = \sum_{p=0}^{\infty}\hat{\mathscr{U}}_I^{(p)}(\beta,0)\,,$$

in which $\hat{\mathscr{U}}_I^{(0)}(\beta,0) = \hat{1}$ and

$$\hat{\mathscr{U}}_I^{(p+1)}(\beta,0) = -\int_0^\beta d\beta'\,\hat{\mathscr{V}}^I_{\rm int}(\beta')\,\hat{\mathscr{U}}_I^{(p)}(\beta',0)\,.$$

Since $e^{-\beta\hat{\mathscr{K}}} = e^{-\beta\hat{\mathscr{K}}_0}\,\hat{\mathscr{U}}_I(\beta,0)$,

$$\begin{aligned}{\rm Tr}_{\mathcal{H}}\Big(e^{-\beta\hat{\mathscr{K}}}\Big) &= {\rm Tr}_{\mathcal{H}}\Big(e^{-\beta\hat{\mathscr{K}}_0}\Big) + \sum_{p=1}^{\infty}{\rm Tr}_{\mathcal{H}}\Big(e^{-\beta\hat{\mathscr{K}}_0}\,\hat{\mathscr{U}}_I^{(p)}(\beta,0)\Big)\\ &= \Xi^{(0)}_{\rm stat}\left\{1 + \sum_{p=1}^{\infty}\langle\hat{\mathscr{U}}_I^{(p)}(\beta,0)\rangle_0\right\},\end{aligned}$$

one obtains in this way the perturbative expansion of the statistical sum $\Xi^{(0)}_{\rm stat}$ of which the first term is the statistical sum of the noninteracting system and the corrections to it take the form of means (denoted $\langle\dots\rangle_0$) with respect to the free statistical operator

$$\hat{\rho}^{(0)} = \frac{1}{\Xi^{(0)}_{\rm stat}}\,e^{-\beta(\hat{\mathscr{H}}_0-\mu\hat{\mathscr{N}})}\,, \tag{358}$$

---

[206] Who attended my Advanced Quantum Mechanics course will easily recognize some known formulae in a new disguise.

which are naturally computed in the basis of the eigenvectors of the free Hamiltonian $\hat{\mathscr{H}}_0$ exploiting the associated distributions (339) and Fermi levels.

Another circumstance owing to which the developed quantum theory of systems consisting of many nointeracting elements is not only of academic interest is the following. In many cases of interacting systems (also strongly interacting ones) it turns out to be possible to introduce in the big Hilbert space (326) an alternative basis which has the structure similar to the Fock space of vectors $|n_1, n_2, \dots\rangle$ generated out of the vector $|\mathrm{void}\rangle$ by the action of the original creation operators and to associate with this new basis another set of creation and annihilation operators which are related in a complicated way to the original ones but satisfy the same commutation relations (the transition from the original to the new creation and annihilation operators is therefore a sort of a canonical transformation). The Hamiltonian of the theory expressed in terms of the new creation and annihilation operators can again be split into $\hat{\mathscr{H}}_0^{\mathrm{new}}$ and $\hat{\mathscr{V}}_{\mathrm{int}}^{\mathrm{new}}$ with $\hat{\mathscr{H}}_0^{\mathrm{new}}$ which has the form of the Hamiltonian (336) of nonineracting elements (called quasiparticles) but the interaction term $\hat{\mathscr{V}}_{\mathrm{int}}^{\mathrm{new}}$ representing the interaction of quasiparticles is now small and can be, in the first approximation, neglected. Therefore the whole quantum statistical theory of noninteracting systems developed in this Lecture can be applied and its results usually reflect quite well the behaviour of real strongly interacting systems which can be interpreted in terms of weakly interacting quasiparticles. Of course the approach based on quasiparticles can be, if their mutual interaction can be obtained (exactly, or in some approximation) from the original Hamiltonian, combined with the perturbative expansion outlined above.

LECTURE XIV (STAT)

In Lecture XIII, using the Grand Canonical Ensemble, we have considered quantum perfect gases of indistinguishable particles. We have obtained two rigorous results: the relation $p\,V = (2/3)U$ and the equation $p\,V^{5/3} = \text{Const.}$ of an adiabatic change which hold true irrespectively of the nature of the particles forming the gas, that is, irrespectively of whether they are bosons or fermions. We have also obtained the leading correction to the perfect gas equation of state in the regime $|z| \ll 1$ in which the Boltzmann approximation becomes reliable. Although the correction does depend on the nature of particles, it too can be obtained essentially by the same steps in both these cases. Other properties of perfect gases of fermions and bosons which are specific for their behavious in the low temperature regime are however markedly different and require separate treatements. We consider the gas of fermion first. The results which will be obtained apply for example to the gas of conduction electrons in metals. Its heat capacity adds up to the heat capacity of the lattice (which can be calculated using the Debye approximation) to give the total heat capacity of a conducting solid at low temperatures.

Zero temperature energy and pressure of the perfect gas of fermions

As a first step towards considering properties the perfect gas of indistinguishable fermions in the low temperature regime we compute the pressure exerted by N fermions of mass m and a half-integer spin s ($g_s = 2s + 1$ internal spin states) enclosed in the box of volume $V = L^3$ at exactly zero temperature. In doing this we will introduce some quantities which will prove useful in further computations.

At $T \to 0$ the mean occupancy of the one-fermion state $|\mathbf{p}, \sigma\rangle$ given by the lower formula (339) takes the form of the simple "step" function of the energy $\varepsilon_{\mathbf{p}}$ of this state:

$$\overline{n_{\mathbf{p},\sigma}} = \frac{1}{1 + \exp((\varepsilon_{\mathbf{p}} - \mu)/k_{\mathrm{B}}T)} \longrightarrow \begin{cases} 1\,, & \text{if } \varepsilon_{\mathbf{p}} < \mu \\ 0\,, & \text{if } \varepsilon_{\mathbf{p}} > \mu \end{cases}. \tag{359}$$

Owing to this, determination of the chemical potential μ of the Grand Canonical Ensemble by fitting it at fixed temperature T and volume V to the (mean, but we have already agreed upon not writing the bar over it) number N of particles using the relation

$$N = g_s V \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\, \overline{n_{\mathbf{p},\sigma}}\,,$$

becomes straightforward at $T = 0$. The zero temperature chemical potential through the equalities

$$\mu(T = 0, v) \equiv \varepsilon_{\mathrm{F}} = \frac{p_{\mathrm{F}}^2}{2m}\,,$$

determines the **Fermi energy** ε_{F} and the corresponding **Fermi momentum** p_{F} at which, because of the form (359) of the Fermi-Dirac distribution, all integrals over momenta are cutted. In turn, these are fixed by the condition

$$\frac{N}{V} \equiv \frac{1}{v} = 4\pi\, \frac{g_s}{(2\pi\hbar)^3} \int_0^{p_{\mathrm{F}}} dp\, p^2 = \frac{4}{3}\,\pi\, \frac{g_s}{(2\pi\hbar)^3}\, p_{\mathrm{F}}^3 = \frac{g_s}{6\pi^2} \left(\frac{p_{\mathrm{F}}}{\hbar}\right)^3.$$

Hence,

$$p_{\rm F} = \hbar\left(\frac{6\pi^2}{g_s}\,\frac{1}{v}\right)^{1/3}, \qquad \varepsilon_{\rm F} = \frac{\hbar^2}{2m}\left(\frac{6\pi^2}{g_s}\,\frac{1}{v}\right)^{2/3}. \tag{360}$$

At exactly zero temperature all single-particle states $|\mathbf{p},\sigma\rangle$ of energies lower than the Fermi energy $\varepsilon_{\rm F}$ ($|\mathbf{p}| < p_{\rm F}$) are occupied ($\overline{n_{\mathbf{p},\sigma}} = 1$), while those of energies greater than $\varepsilon_{\rm F}$ ($|\mathbf{p}| > p_{\rm F}$) are empty ($\overline{n_{\mathbf{p},\sigma}} = 0$). Electrons are said to fill in the momentum space the volume enclosed by the **Fermi shell** defined by the condition $|\mathbf{p}| = p_{\rm F}$. The internal energy U (which is simply the energy of the ground state of the system of noninteracting fermions) can now be easily computed as

$$U = g_s V\,\frac{4\pi}{(2\pi\hbar)^3}\int_0^{p_{\rm F}} dp\,p^2\,\frac{p^2}{2m} = \frac{3}{5}\,N\,\frac{p_{\rm F}^2}{2m} = \frac{3}{5}\,N\,\varepsilon_{\rm F}\,. \tag{361}$$

The pressure of the (quantum) gas of fermions at $T = 0$ (one speaks of the **degenerate gas of fermions**) follows now directly from the first one of the three formulae (342):

$$p = \lim_{T\to 0} g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\,k_{\rm B}T\ln\!\left(1 + e^{-(\varepsilon_{\mathbf{p}}-\mu)/k_{\rm B}T}\right).$$

The contribution to the integral from the region $\varepsilon_{\mathbf{p}} > \mu$ vanishes in the indicated limit because both the logarithm itself and the factor $k_{\rm B}T$ tend to zero. In turn, in the region in which $\varepsilon_{\mathbf{p}} < \mu$ the exponens under the logarithm diverges, the factor $+1$ can be neglected and (replacing $\mu(T = 0, v)$ by $\varepsilon_{\rm F}$ and hoping that the reader is smart enough not to confuse the integration variable p with the pressure p) one obtains

$$p = g_s\,\frac{4\pi}{(2\pi\hbar)^3}\int_0^{p_{\rm F}} dp\,p^2\left(\varepsilon_{\rm F} - \frac{p^2}{2m}\right) = g_s\,\frac{4\pi}{(2\pi\hbar)^3}\,p_{\rm F}^3\left(\frac{\varepsilon_{\rm F}}{3} - \frac{p_{\rm F}^2}{10m}\right).$$

The zerotemperature pressure of the quantum gas of fermions which upon using the result (360) can be written in the form

$$p = \frac{g_s}{30\pi^2}\,\frac{\hbar^2}{m}\left(\frac{6\pi^2}{g_s}\,\frac{1}{v}\right)^{5/3},$$

is that pressure which, as it has first been understood by S. Chandrasekhar, opposing itself to gravity, maintains the equilibrium of neutron stars.[207]

The Sommerfeld expansion

We now set ourselves to determine the properties of the perfect gas of fermions. We will be interested in particular in the behaviour of its entropy and heat capacity in the limit $T \to 0$. To this end we must analyse the formulae (342) in the regime $k_{\rm B}T \ll \varepsilon_{\rm F}$, that is in the limit $z \gg 1$ (opposite to the one in which the Boltzmann approximation works).

---

[207] In fact in deriving the Chandrasekhar result one should use the relativistic formula for the energy of particles.

The technical mean to achieve this goal is the so-called Sommerfeld expansion which allows to compute in the limit $\beta \to \infty$ integrals of the general form

$$\int_0^\infty d\varepsilon\, f(\varepsilon)\,\frac{1}{1+e^{\beta(\varepsilon-\mu)}}. \tag{362}$$

In this limit the factor $[1+\exp(\beta(\varepsilon-\mu)]^{-1}$ under the integral is very close to the "step" function considered above and its derivative with respect to ε is nonvanishing practically only in the vicinity of the point $\varepsilon = \varepsilon_{\rm F}$, say in the domain $|\varepsilon - \varepsilon_{\rm F}| \sim k_{\rm B}T$. This suggests integration of (362) by parts[208] which will precisely yield under the integral the derivative of the factor $[1+\exp(\beta(\varepsilon-\mu)]^{-1}$; it will be then possible to expand the remaining factors of the integrand in the Taylor series around $\varepsilon = \varepsilon_{\rm F}$.

To stress the general validity of the final formula[209] we will deal with the integral (362) written in the "neutral" variable x. Let $F(x)$ be the primitive function of $f(x)$, that is, such that $F'(x) = f(x)$. Integrating (362) by parts we obtain

$$\int_0^\infty dx\,\frac{f(x)}{1+e^{x-\nu}} = -\frac{F(0)}{1+e^{-\nu}} + \int_0^\infty dx\,F(x)\,\frac{e^{x-\nu}}{(1+e^{x-\nu})^2}\,.$$

It has been assumed here that $F(x)\,e^{-x} \to 0$ as $x \to \infty$. Furthermore, the primitive F of the function f can alway be defined so that $F(0) = 0$. With such a definition the whole boundary term vanishes. The function multiplying $F(x)$ under the integral on the right hand side vanishes very fast, essentially like $e^{-|x-\nu|}$, both to left and to the right of the point $x = \nu$. Therefore expanding $F(x)$ in the Taylor series around this point and integrating term by term the resulting integrand should yield a reasonable approximation to the original integral. After shifting the integration variable, $\xi = x - \nu$, one obtains in this way

$$\int_0^\infty dx\,\frac{f(x)}{1+e^{x-\nu}} = \int_{-\nu}^\infty d\xi\left[F(\nu) + \xi\,f(\nu) + \frac{1}{2}\,\xi^2 f'(\nu) + \frac{1}{6}\,\xi^3 f''(\nu) + \dots\right]\frac{e^\xi}{(1+e^\xi)^2}\,.$$

If $\nu \gg 1$, which is precisely the case when the integrals (362) are considered in the limit $T \to 0$, extending the lower integration limit to $-\infty$ should result in an only exponentially small, i.e. of order $e^{-\nu}$, error.[210]

Thus, to obtain the series expansion of the original integral one needs to evaluate the integrals

$$I_n = \int_{-\infty}^\infty d\xi\,\xi^n\,\frac{e^\xi}{(1+e^\xi)^2}\,.$$

---

[208] The Sommerfelda expansion can be also derived differently - see e.g. the textbook of Schwabl.

[209] Its derivation involves some fancy mathematics which can be omitted, if the reader doesn't like it; what matters is only the result (363).

[210] Nevertheless, this operation has the effect that the resulting series representing the original (decent, i.e. perfectly convergent) integral is not convergent, but only asymptotic. This follows from the fact that the similar extension of the lower integration limit done in the original integral produces, if the function $f(x)$ does not tend to zero fast enough as $x \to -\infty$ (and we typically will consider functions $f(x) = x^\alpha$ with a fractional exponent α, which require in addition to be further specified for $x < 0$), a divergent integral.

The zeroth one, I_0, is elementary:[211]

$$I_0 = \int_{-\infty}^{\infty} d\xi \, \frac{e^{\xi}}{(1+e^{\xi})^2} = -\int_{-\infty}^{\infty} d\xi \, \frac{d}{d\xi} \, \frac{1}{1+e^{\xi}} = 1 \, .$$

Next, making the substitiution $\xi = -\eta$, one notices that

$$I_n = -\int_{\infty}^{-\infty} d\eta \, (-\eta)^n \, \frac{e^{-\eta}}{(1+e^{-\eta})^2} = (-1)^n \int_{-\infty}^{\infty} d\eta \, \eta^n \, \frac{e^{\eta}}{(1+e^{\eta})^2} = (-1)^n I_n \, .$$

It follows that $I_{2n+1} = 0$, while the integrals I_{2n} (their integrands are even functions of ξ) can be taken from 0 to ∞ if the result is doubled; these integrals[212] can be therefore written in the form

$$I_{2n} = 2\left[-\frac{d}{d\lambda}\int_0^{\infty} d\xi \, \frac{\xi^{2n-1}}{1+e^{\lambda\xi}}\right]_{\lambda=1} = 2\left[-\frac{d}{d\lambda}\,\frac{1}{\lambda^{2n}}\right]_{\lambda=1}\int_0^{\infty} d\eta \, \frac{\eta^{2n-1}}{1+e^{\eta}} = 4n\int_0^{\infty} d\eta \, \frac{\eta^{2n-1}}{1+e^{\eta}} \, .$$

(In the second step the integration variable has been rescaled.) One is therefore led to consider the integrals

$$J_p = \int_0^{\infty} d\xi \, \frac{\xi^{p-1}}{1+e^{\xi}} = \Gamma(p) f_p(1) \, ,$$

The functions $f_p(z)$, defined in (349), can be expressed as infinite series (350) which are convergent at $z = 1$ if $p > 1$. Their values $f_p(1)$ can be expressed through the famous Riemann ζ-function. To this end one writes two different infinite power series (absolutely convergent if $p > 1$) splitting each of them into two sums:[213]

$$\begin{aligned} \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l^p} &= -\sum_{l=1}^{\infty} \frac{1}{(2l)^p} + \sum_{l=0}^{\infty} \frac{1}{(2l+1)^p} \equiv f_p(1) \, , \\ \sum_{l=1}^{\infty} \frac{1}{l^p} &= \sum_{l=1}^{\infty} \frac{1}{(2l)^p} + \sum_{l=0}^{\infty} \frac{1}{(2l+1)^p} \equiv b_p(1) = \zeta(p) \, . \end{aligned}$$

Subtracting them side by side one obtains the relation

$$f_p(1) - b_p(1) = -2\sum_{l=1}^{\infty} \frac{1}{(2l)^p} = -2^{1-p} \, b_p(1) \, ,$$

---

[211] I_0 can be also evaluated in the limits $[-\nu, \infty)$. This allows to check that the expansion does not depend on the adopted assumption that $F(0) = 0$; moreover, if $F(0) \neq 0$ and the lower limit of the integral is extended to $-\infty$ the boundary term would be not canceled but if $\nu \gg 1$ it would be $\sim F(0)\, e^{-\nu}$, that is of the same order as the error introduced by extending the lower integration limit to $-\infty$.

[212] Except for I_0.

[213] The second series is denoted $b_p(1)$, because it gives the values at $z = 1$ of the functions $b_p(z)$ - the "bosonic" analogs of the functions $f_p(z)$ - which play an important role in analysing low temperature properties of noninteracting bosons.

which implies, because $b_p(1) = \zeta(p)$, that

$$f_p(1) = \left(1 - 2^{1-p}\right) b_p(1) = \left(1 - 2^{1-p}\right) \zeta(p) \,.$$

Therefore,

$$I_{2n} = 4n \, J_{2n} = 4n \, (1 - 2^{1-2n}) \, \zeta(2n) \, \Gamma(2n) \,.$$

The values $\zeta(2n)$ of the Riemann ζ-function can be obtained by summing infinite series using e.g. methods of the complex analysis (see e.g. my notes to complex integrals) which yield[214]

$$\zeta(2) = \frac{\pi^2}{6} \,, \qquad \zeta(4) = \frac{\pi^4}{90} \,, \qquad \zeta(6) = \frac{\pi^6}{945} \,, \qquad \dots$$

Using these mathematical results we can finally write the Sommerfeld expansion[215]

$$\int_0^\infty dx \, \frac{f(x)}{1 + e^{x-\nu}} = F(\nu) + \frac{\pi^2}{6} \, f'(\nu) + \frac{7\pi^4}{360} \, f'''(\nu) + \frac{31\pi^6}{15120} \, f^{(5)}(\nu) + \dots \tag{363}$$

Gas of noninteracting fermions in the low temperature regime
We can now analyse the content of the formulae (342), expressed below in terms of the functions $f_{5/2}(z)$ and $f_{3/2}(z)$,

$$\begin{aligned} \frac{p}{k_{\rm B}T} &= g_s \left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} f_{5/2}(z) \,, \\ \frac{1}{v} &= g_s \left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} f_{3/2}(z) \,, \\ \frac{U}{V} &= \frac{3}{2} \, k_{\rm B}T \, g_s \left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} f_{5/2}(z) = \frac{3}{2} \, p \,, \end{aligned} \tag{364}$$

in the regime $z \gg 1$. We first apply the Sommerfeld expansion (363) to the functions $f_{3/2}(z)$ i $f_{5/2}(z)$: taking into account their "nonperturbative" definition (349) we find (recall that $\nu = \ln z = \mu/k_{\rm B}T$):

$$f_{3/2}(z) = \frac{2}{\sqrt{\pi}} \int_0^\infty dx \, \frac{x^{1/2}}{1 + e^{x-\nu}} = \frac{2}{\sqrt{\pi}} \left\{ \frac{2}{3} \, \nu^{3/2} + \frac{\pi^2}{12} \, \nu^{-1/2} + \frac{7\pi^4}{960} \, \nu^{-5/2} + \dots \right\} , \tag{365}$$

---

[214] We note that the integrals J_{2n} can be also written in terms of the Bernoulli numbers B_n

$$J_{2n} = \int_0^\eta d\eta \, \frac{\eta^{2n-1}}{1 + e^\eta} = \frac{2^{2n-1} - 1}{2n} \, \pi^{2n} \, B_n \,,$$

where $B_1 = 1/6$, $B_2 = 1/30$, $B_3 = 1/42$, $B_4 = 1/6$, etc.

[215] This is the final formula, which is important; recall that $F(x)$ is the primitive function of $f(x)$ satisfying the condition $F(0) = 0$.

$$\frac{3}{2}\,f_{5/2}(z)=\frac{2}{\sqrt{\pi}}\int_0^\infty dx\,\frac{x^{3/2}}{1+e^{x-\nu}}=\frac{2}{\sqrt{\pi}}\left\{\frac{2}{5}\,\nu^{5/2}+\frac{\pi^2}{4}\,\nu^{1/2}-\frac{7\pi^4}{960}\,\nu^{-3/2}+\dots\right\}. \qquad (366)$$

The formula for $1/v$ can be, therefore, (extracting the factors of $2/3$ and $\nu^{3/2}\equiv(\mu/k_{\rm B}T)^{3/2}$ from the curly bracket and combining the latter with the bracket depending on the fermion's mass) written in the form

$$\frac{1}{v}=\frac{4g_s}{3\sqrt{\pi}}\left(\frac{m\mu}{2\pi\hbar^2}\right)^{3/2}\left\{1+\frac{\pi^2}{8}\,\frac{1}{\nu^2}+\frac{7\pi^4}{640}\,\frac{1}{\nu^4}+\dots\right\}.$$

Multiplying both sides of this relation by v expressed in the right hand side in terms of the Fermi energy $\varepsilon_{\rm F}$ defined in (360):

$$v=\frac{6\pi^2}{g_s}\left(\frac{\hbar^2}{2m\varepsilon_{\rm F}}\right)^{3/2},$$

we obtain the identity

$$1=\left(\frac{\mu}{\varepsilon_{\rm F}}\right)^{3/2}\left\{1+\frac{\pi^2}{8}\left(\frac{k_{\rm B}T}{\mu}\right)^2+\frac{7\pi^4}{640}\left(\frac{k_{\rm B}T}{\mu}\right)^4+\dots\right\},$$

which determines the chemical potential μ as a function of T and $v=V/N$ (the dependence on this quantity is hidden in $\varepsilon_{\rm F}$). As we are interested in the limit $k_{\rm B}T\ll\varepsilon_{\rm F}\equiv\mu(0,v)$, we insert in the right hand side of this identity[216]

$$\mu=\varepsilon_{\rm F}(1+x)\,,$$

which, after expanding in powers of x gives the relation

$$1=\left(1+\frac{3}{2}\,x+\dots\right)\left[1+\frac{\pi^2}{8}\left(\frac{T}{T_{\rm F}}\right)^2-\frac{\pi^2}{4}\left(\frac{T}{T_{\rm F}}\right)^2x+\dots\right],$$

in which $T_{\rm F}\equiv\varepsilon_{\rm F}/k_{\rm B}$. Since x is itself of order $(T/T_{\rm F})^2$, the third and all other terms in the bracket can be discarded and one obtains

$$x=-\frac{\pi^2}{12}\left(\frac{T}{T_{\rm F}}\right)^2+\dots$$

that is (the dependence on v is hidden in $\varepsilon_{\rm F}$)

$$\mu(T,v)=\varepsilon_{\rm F}\left[1-\frac{\pi^2}{12}\left(\frac{T}{T_{\rm F}}\right)^2+\mathcal{O}\left(\frac{T^4}{T_{\rm F}^4}\right)\right].$$

---

[216] Actually, it would be more systematic to insert there the Taylor expansion of the chemical potential around $T=0$: $\mu=\varepsilon_{\rm F}(1+a_1x+a_2x^2+\dots)$ with $x\equiv T/T_{\rm F}$ and to determine the coefficients a_i by requiring that the coefficients of all the powers of T (i.e. of x) vanish, but since we want to find only the first term of the expansion, the adopted method is sufficient.

Knowing the chemical potential as a function of T and $v = V/N$ we can now express the internal energy and entropy through T, V, and N and compute the heat capacity of the quantum gas of noninteracting fermions. To this end, using the expansion of the function $(3/2)f_{5/2}(z)$, we write (putting the factor $(k_{\rm B}T)^{5/2}$ into the curly bracket and replacing ν by $\mu/k_{\rm B}T$)

$$\frac{U}{V} = g_s\left(\frac{2m}{\hbar^2}\right)^{3/2}\left(\frac{1}{4\pi}\right)^{3/2}\frac{4}{5\sqrt{\pi}}\left\{\mu^{5/2} + \frac{5}{8}\pi^2(k_{\rm B}T)^2\mu^{1/2} + \dots\right\}. \tag{367}$$

Inserting here (cf. the definition (360))

$$\frac{2m}{\hbar^2} = \frac{1}{\varepsilon_{\rm F}}\left(\frac{6\pi^2}{g_s}\frac{1}{v}\right)^{2/3},$$

and $\mu = \varepsilon_{\rm F}(1 + x)$, we obtain

$$\frac{U}{V} = \frac{3}{5}\frac{\varepsilon_{\rm F}}{v}\left[\left(1 + \frac{5}{2}x + \dots\right) + \frac{5}{8}\pi^2\left(\frac{T}{T_{\rm F}}\right)^2\left(1 + \frac{1}{2}x + \dots\right)\right],$$

or, because $1/v = N/V$ (we remember that the dependence on v is also hidden in $\varepsilon_{\rm F}$),

$$U = \frac{3}{5}N\varepsilon_{\rm F}\left(1 + \frac{5}{12}\pi^2\left(\frac{T}{T_{\rm F}}\right)^2 + \dots\right). \tag{368}$$

This of course agrees with the result (361). The heat capacity C_V can be now easily computed:

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{V,N} = \frac{\pi^2}{2}N\varepsilon_{\rm F}\frac{T}{T_{\rm F}^2} + \dots = \frac{\pi^2}{2}Nk_{\rm B}\frac{T}{T_{\rm F}} + \dots \tag{369}$$

It vanishes as $T \to 0$, as required by 3TMDL.

In turn the entropy of the quantum gas of noninteracting fermions can be obtained as the derivative

$$S = -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} = \left(\frac{\partial (pV)}{\partial T}\right)_{V,\mu} = \frac{2}{3}\left(\frac{\partial U}{\partial T}\right)_{V,\mu}.$$

It should be noted that this is the derivative at constant μ (not at constant N, as in computing C_V). It should, therefore, be taken before the chemical potential μ in the formula (367) for U is expressed through T and v:

$$S = \frac{2}{3}Vg_s\left(\frac{2m}{\hbar^2}\right)^{3/2}\left(\frac{1}{4\pi}\right)^{3/2}\frac{4}{5\sqrt{\pi}}\frac{\partial}{\partial T}\left\{\mu^{5/2} + \frac{5}{8}\pi^2(k_{\rm B}T)^2\mu^{1/2} + \dots\right\}.$$

Only after the derivative is taken can one substitute here $\mu = \varepsilon_{\rm F} + \dots$ to get

$$S = \frac{\pi^2}{2}Nk_{\rm B}\frac{T}{T_{\rm F}} + \dots$$

As required by 3TMDL, entropy vanishes in the limit $T \to 0$. And, of course, the heat capacity computed using the formula $C_V = T(\partial S/\partial T)_{V,N}$ gives again the result (369).

The equation of state of the gas of fermions valid in the limit $k_{\rm B}T \ll \varepsilon_{\rm F}$ can be obtained from the rigorously proved relation (344) expressing in it U in terms of T using (368):

$$pV = \frac{2}{5}\,N\varepsilon_{\rm F}\left(1 + \frac{5}{12}\pi^2\left(\frac{T}{T_{\rm F}}\right)^2 + \dots\right).$$

Entropy can be now obtained in the alternative way, without differentiation (therefore the chemical potential can from the beginning be expressed throught T and v) using the relation $\Omega = U - TS - \mu N$:

$$\frac{S}{Nk_{\rm B}} = \frac{U - \Omega - \mu N}{Nk_{\rm B}T} = \frac{5}{2}\,\frac{pV}{Nk_{\rm B}T} - \frac{\mu}{k_{\rm B}T}\,.$$

The second form follows from the fact that $\Omega = -pV$ and the (rigorous) relation (344) $U = (3/2)pV$. Therefore,

$$\frac{S}{Nk_{\rm B}} = \frac{5}{2}\,\frac{p\,v}{k_{\rm B}T} - \nu\,.$$

Expressing here the ratio $p\,v/k_{\rm B}T$ as in (351) throught the ratio of the $f_{5/2}$ and $f_{3/2}$ functions and using their Sommerfeld expansions[217] (366) and (365) we get

$$\frac{S}{Nk_{\rm B}} = \frac{5}{2}\,\frac{f_{5/2}(\nu)}{f_{3/2}(\nu)} - \nu = \nu\,\frac{1 + (5/8)\pi^2\nu^{-2} + \dots}{1 + (1/8)\pi^2\nu^{-2} + \dots} - \nu = \frac{\pi^2}{2}\,\frac{1}{\nu} + \dots = \frac{\pi^2}{2}\,\frac{k_{\rm B}T}{\mu} + \dots$$

To this order, however, $\mu = \varepsilon_{\rm F} = k_{\rm B}T_{\rm F}$, so this is the result obtained previously.

Quantum perfect gas of bosons

Analysing low temperature properties of the quantum gas of noninteracting indistinguishable fermions we have learned that at exactly zero temperature such particles occupy all single-particle states of (single-particle) energies lower than the Fermi energy $\varepsilon_{\rm F}$ which is the same as the zero temperature chemical potential and is determined by the gas density $1/v$ (and the particle's halfinteger spin) and that the pressure of the gas is not zero - a fact which is crucial for final stages of evolution of stars. We have also seen that, as required by 3TMDL, entropy S and heat capacity C_V of the gas vanish at $T = 0$ (both like $\sim$ Const. T).

We now consider properties of the gas of noninteracting indistinguishable bosons in the analogous regime. It will turn out that although entropy and heat capacity of the gas of bosons, similarly as those of the gas of fermions, vanish at $T = 0$ (again in agreement with 3TMDL), other its properties are markedly different due to symmetry (instead of

---

[217] Note that we expand now $f_{5/2}(z)$ and not $(3/2)f_{5/2}(z)$, as we did previously.

antisymmetry) of state vectors representing systems of indistinguishable bosons. In particular at sufficiently low temperature in the gas of bosons the so-called Bose-Einstein condensation occurs.

We begin by recalling once more the three general formulae (now specified to bosons) which encode the complete thermodynamical information about the system:

$$\begin{aligned}
&\Omega = -p\,V = -k_{\rm B}T\ln\Xi_{\rm stat} = k_{\rm B}T\sum_l \ln\big(1 - z\,e^{-\beta\varepsilon_l}\big)\,,\\
&N = \sum_l \frac{1}{z^{-1}e^{\beta\varepsilon_l}-1}\,,\\
&U = \sum_l \frac{\varepsilon_l}{z^{-1}e^{\beta\varepsilon_l}-1}\,.
\end{aligned}\tag{370}$$

The sums are over single-particle states $|l\rangle$ which are eigenstates of the single-particle Hamiltonian $\hat{\mathscr{H}}^{(1)}$ with energies ε_l and $z = \exp(\beta\mu)$ is the activity. We recall also that the convergence of the summations over the occupation numbers of these single-particle states, which gave rise to these formulae, required (unlike the case of fermions) that the chemical potential μ is less than the lowest one of the energies ε_l.

If the gas of noninteracting nonrelativistic bosons of integer spin s ($g_s = 2s+1$ spin states) is enclosed in the box of volume $V = L^3$ and periodic boundary conditions are imposed, the role of the states $|l\rangle$ is played by the states $|\mathbf{p},\sigma\rangle$ with momenta $\mathbf{p} = (2\pi/L)\mathbf{n}$ and $\sigma = -s,\dots,+s$. The corresponding energies are $\varepsilon_{\mathbf{p},\sigma} \equiv \varepsilon_{\mathbf{p}} = \mathbf{p}^2/2m$. The summation over the labels l in the formulae (370) runs in this case over all triplets $\mathbf{n} = (n_x, n_y, n_z)$ of integers and over σ. As usually, when the system is large (macroscopic), and as was done in the case of the gas of fermions, one wants to make the standard replacement

$$\sum_{\sigma=-s}^{s}\sum_{\mathbf{p}} \to g_sV\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\,.\tag{371}$$

One has however be careful in this case and think about the physics of the system. As $T\to 0$, all bosons can and certainly will (just because they like to do so!) occupy the single one-particle state corresponding to $\mathbf{p} = \mathbf{0}$ that is to $\varepsilon_l = 0$. If the summation over l is replaced by the integration according to the prescription (371), the contribution of this state will be lost (because the measure $dp\,p^2$ vanishes at $p = |\mathbf{p}| = 0$). Therefore one has to take it (and perhaps also the contribution of the other states corresponding to $|\mathbf{p}|\sim 0$) into account separately by writing the formulae (370) in the forms

$$\begin{aligned}
\frac{p}{k_{\rm B}T} &= -g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\ln\big(1 - z\,e^{-\varepsilon_{\mathbf{p}}/k_{\rm B}T}\big) - \frac{g_s}{V}\ln(1-z) - \frac{g_s}{V}\ln\big(1 - z\,e^{-\varepsilon_1/k_{\rm B}T}\big) + \dots,\\
\frac{1}{v} &= g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\frac{1}{z^{-1}e^{\varepsilon_{\mathbf{p}}/k_{\rm B}T}-1} + \frac{g_s}{V}\frac{z}{1-z} + \frac{g_s}{V}\frac{1}{z^{-1}e^{\varepsilon_1/k_{\rm B}T}-1} + \dots,\\
\frac{U}{V} &= g_s\int\frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\frac{\varepsilon_{\mathbf{p}}}{z^{-1}e^{\varepsilon_{\mathbf{p}}/k_{\rm B}T}-1} + 0 + \frac{g_s}{V}\frac{\varepsilon_1}{z^{-1}e^{\varepsilon_1/k_{\rm B}T}-1} + \dots
\end{aligned}\tag{372}$$

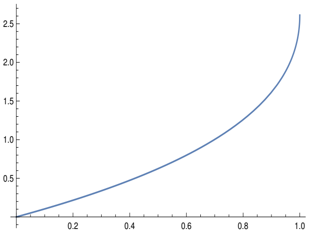

Figure 23: Plot of the function $b_{3/2}(z)$. Its value at $z = 1$ is 2.612.

in which the contributions of the one-particle states of energies $\varepsilon_0 = 0$, ε_1, ..., have been explicitly singled out.

To proceed, it is convenient to introduce the family of functions $b_p(z)$, the bosonic counterparts of the functions $f_p(z)$ (cf. the formulae (349) and (350)) defined by

$$b_p(z) = \frac{1}{\Gamma(p)} \int_0^\infty dx\, x^{p-1} \frac{1}{z^{-1}e^x - 1} = \sum_{l=1}^{\infty} \frac{z^l}{l^p}\,. \tag{373}$$

Their series expansions, convergent because $|z| < 1$, as a result of the condition $\mu < \varepsilon_0 = 0$, is obtained in the same way as the expansions of the $f_p(z)$ functions (see Lecture XIII). Written in terms of these functions the equations (372) take the form[218]

$$\begin{aligned}
\frac{p}{k_{\rm B}T} &= \frac{g_s}{\lambda_T^3}\, b_{5/2}(z) - \frac{g_s}{V}\, \ln(1-z) - \frac{g_s}{V}\, \ln\!\left(1 - z\, e^{-\varepsilon_1/k_{\rm B}T}\right) + \dots\,, \\
\frac{1}{v} &= \frac{g_s}{\lambda_T^3}\, b_{3/2}(z) + \frac{g_s}{V}\, \frac{z}{1-z} + \frac{g_s}{V}\, \frac{1}{z^{-1}e^{\varepsilon_1/k_{\rm B}T} - 1} + \dots\,, \\
\frac{U}{V} &= \frac{3}{2}\, k_{\rm B}T\, \frac{g_s}{\lambda_T^3}\, b_{5/2}(z) + 0 + \frac{g_s}{V}\, \frac{\varepsilon_1}{z^{-1}e^{\varepsilon_1/k_{\rm B}T} - 1} + \dots
\end{aligned} \tag{374}$$

$\lambda_T = (2\pi\hbar/mk_{\rm B}T)^{1/2}$ is here the thermal wavelength of the particle.

As in the case of fermions, the first step towards the investigation of the gas properties is to determine the chemical potential μ or the activity z as a function of T and v from the second one of the relations (374). It is convenient to rewrite it in the form

$$\frac{\lambda_T^3}{g_s v} = b_{3/2}(z) + \left(\frac{\lambda_T^3}{g_s v}\right) \frac{g_s}{N}\, \frac{z}{1-z} + \left(\frac{\lambda_T^3}{g_s v}\right) \frac{g_s}{N}\, \frac{1}{z^{-1}e^{\varepsilon_1/k_{\rm B}T} - 1} + \dots\,. \tag{375}$$

The dependence on z of the first term on the right hand side is shown in Figure 23. What is important for the analysis below is that the function $b_{3/2}(z)$ is monotonically increasing and bounded in the domain $0 \leq z \leq 1$, reaching at $z = 1$ its maximal value equal $\zeta(3/2) = 2.612$; this value is obviously independent of N. When the left hand side of the equation (375) increases (either because $T \to 0$ or because the gas density $\rho = 1/v$

---

[218] The first of these equations is written in terms of the function $b_{5/2}(z)$ after integration by parts of its right hand side; in writing the third one the rigorously proved (in Lecture XIII) relation $pV = (2/3)U$ - but applied here only to the terms given by integrals - has been used.

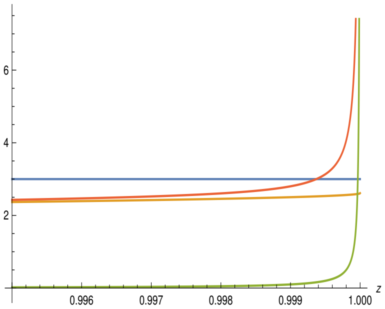

Figure 24: Graphical determination of the activity z. $N = 10^4$, $\lambda_T^3/v = 3$, $g_s = 1$. The four curves (going from below near the vertical axis) show the values of: a) $(z/(1-z))/N$, b) $b_{3/2}(z)$, c) $b_{3/2}(z) + (\lambda_T^3/V)\,z\,(1-z)^{-1}$, d) λ_T^3/v. Owing to the contribution of the term $(\lambda_T^3/V)\,z\,(1-z)^{-1}$ which for finite N (finite V) diverges at $z = 1$, the solution exists even if $\lambda_T^3/v > \zeta(3/2)$. As $N \to \infty$ with v fixed, the deviation of the solution z from unity is $\sim 1/N$.

increases) but remains smaller that $\zeta(3/2)$, it is balanced essentially by the value of the function $b_{3/2}(z)$ on the right hand side by appropriately increasing the value of z towards unity, while the contribution of the remaining terms is killed in the thermodynamic limit ($N \to \infty$, $V \to \infty$ with V/N kept fixed) because, as is clear from Figure 23, the solution of (375) exists also in the absence of these terms. However, when the factor $\lambda_T^3/g_s v$ exceeds $\zeta(3/2)$, it is precisely the additional terms on the right hand side of (375) which must complement the value of the function $b_{3/2}(z)$ to balance the left hand side; graphical solving the equation (375) for z in this situation at finite value of N is shown in Figure 24. Moreover, of the additional terms present on the right hand side of (375) only the first one is singular at $z = 1$ and therefore it is only this term which can contribute nonnegligibly in the thermodynamic limit:[219]

$$\lim_{\infty}\left(\frac{g_s}{N}\,\frac{z}{1-z}\right) \equiv \lim_{\infty}\frac{\overline{n}_{\mathbf{p}=\mathbf{0}}}{N} = \left(1 - \frac{g_s v}{\lambda_T^3}\,b_{3/2}(1)\right) \neq 0\,. \tag{376}$$

The last equality follows from the fact, clear from Figure 24, that if $N \gg 1$, the deviation of the value of z obtained from (375) from unity is of order $1/N$, so in the strict thermodynamic limit $z = 1$.

It follows that if $\lambda_T^3/g_s v > \zeta(3/2)$, the fraction of the total number of particles occupying the zero-momentum, zero-energy one-particle state is, in the thermodynamic limit, finite. This is the phenomenon called **Bose-Einstein condensation**. The domain in the

---

[219] It becomes now clear that the terms singled out in the third formula of (374), the one for the system's internal energy U, being nonsingular at $z = 1$, simply drop out in the thermodynamic limit. Similarly drop also the corresponding terms in the first formula of (374); the first one because (376) shows that $(g_s/N) \sim (1-z)/z$ and $\lim_{z\to 1}(1-z)\ln(1-z) = 0$, and the remaining ones, because they are explicitly nonsingular at $z = 1$. Therefore the relation (344) derived in Lecture XIII remains true also in the presence of the condensation; heuristically, because the particles in the condensate contribute neither to the total energy of the gas nor to its pressure.

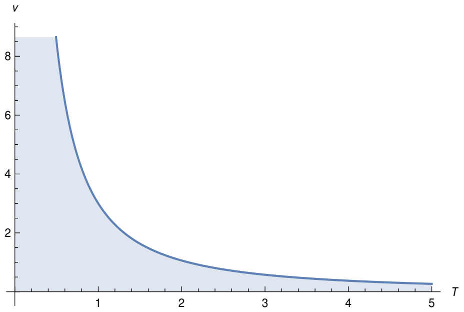

Figure 25: The (T, v) plane of the system of noninteracting identical bosons (arbitrary units). Marked is the region in which the Bose-Einstein condensation occurs, that is one in which the "gaseous" phase coexists with the "liquid-like" phase.

plane (T, v) in which the system of free bosons condenses in this limit is bounded by the curve

$$g_s v \left(\frac{m k_{\rm B} T}{2\pi\hbar^2}\right)^{3/2} \zeta(3/2) = \text{Const.}\, v\, T^{3/2} = 1\,. \tag{377}$$

In the thermodynamical limit the activity is, therefore, given by

$$z = \exp(\mu/k_{\rm B}T) = \begin{cases} b^{-1}_{3/2}(\lambda^3_T/g_s v) & \text{outside the condensation domain} \\ 1 & \text{inside the condensation domain} \end{cases}.$$

The fraction of particles occupying the zero-energy, zero-momentu state is

$$\lim_{\infty} \frac{\overline{n}_{\mathbf{p=0}}}{N} = \begin{cases} 0 & \text{outside the condensation domain} \\ 1 - (g_s v/\lambda^3_T)\zeta(3/2) & \text{inside the condensation domain} \end{cases}.$$

In other words, at fixed gas density $\rho = 1/v$, the condensation appears below the critical temperature $T_{\rm cr}$ determined by the condition (377) and

$$\lim_{\infty} \frac{\overline{n}_{\mathbf{p=0}}}{N} = 1 - \left(\frac{T}{T_{\rm cr}}\right)^{3/2},$$

while at fixed temperature the condensation appears above the critical density and

$$\lim_{\infty} \frac{\overline{n}_{\mathbf{p=0}}}{N} = 1 - \frac{v}{v_{\rm cr}} = 1 - \frac{\rho_{\rm cr}}{\rho}\,.$$

On the (T, v) plane there are therefore two domains (see Figure 25): one in which only the "gaseous" phase of the system exists (in the thermodynamic limit $\overline{n_{\mathbf{p=0}}} = 0$) and another one in which the condensate of bosons, $\overline{n_{\mathbf{p=0}}} \neq 0$, forms and coexists with the gaseous phase. Ultimately, at $T = 0$ or $\rho = \infty$, all particles, being bosons, will occupy the single state (g_s degenerate states, if $s > 0$). The forms of the thermodynamic potentials and quantities characterizing the system are different in these two domains because in the first one the chemical potential, i.e. z, varies with T and v, while in the second one it is constant ($z \equiv 1$ in the thermodynamic limit). This has the consequence that in the condensation region (marked in Figure 25) the pressure p given by the formula

$$p = k_{\rm B}T \, \frac{g_s}{\lambda_T^3} \, b_{5/2}(z) \propto (k_{\rm B}T)^{5/2} \, b_{5/2}(z) \, , \tag{378}$$

is independent of the density $\rho = 1/v$ because in this region $z \equiv 1$. The situation is therefore very similar to the ordinary first order vapour-liquid transition: particles occupying the $\mathbf{p} = \mathbf{0}$ state do not exert any pressure and behave like a liquid phase; increasing the system's density at fixed temperature instead of increasing its pressure causes only more particles to occupy the $\mathbf{p} = \mathbf{0}$ state; as a result isotherms of the system plotted in the (v, p) (or the (ρ, p) plane) have flat parts corresponding to the condensation region. The equation of the curve separating on the (v, p) plane the "gaseous" (vapour) region from the one in which the condensation begins can be obtained by determining $k_{\rm B}T_{\rm cr}$ as a function of v from (377) and inserting it into the formula (378) evaluated at $z = 1$. This gives

$$p \, v^{5/3} = \frac{2\pi\hbar^2}{m} \, \frac{\zeta(5/2)}{[\zeta(3/2)]^{5/3}} \, g_s^{-2/3} \, . \tag{379}$$

On the (v, p) plane the reqion of the phase coexistence is delimited by one of the system's adiabats! The phase diagram of the system of free bosons in the (T, p) coordinates is shown in Figure 26.

Using the Clapeyron-Clausius equation (192) one can also determine the latent heat $q_{\rm cond\to gas}$ (per particle) of the Bose-Einstein condensate ("liquid") - gas transition provided one assumes that the volume per particle of the "liquid" phase, i.e. of the condensate, is zero. The derivative dp/dT is easily determined from the relation (378) evaluated at $z = 1$. Thus

$$q_{\rm cond\to gas} = \frac{5}{2} \, k_{\rm B} \, g_s \left(\frac{m k_{\rm B} T}{2\pi\hbar^2} \right)^{3/2} \zeta(5/2) \, T \, v_{\rm cr}(T) \, .$$

Iserting here $v_{\rm cr}(T)$ determined from the relation (377) one obtains the simple result

$$q_{\rm cond\to gas} = \frac{5}{2} \, k_{\rm B}T \, \frac{\zeta(5/2)}{\zeta(3/2)} \, . \tag{380}$$

We now determine other thermodynamic functions characterizing the system of nointeracting identical bosons. The internal energy per particle is given by

$$\frac{U}{N} = \begin{cases} (3/2)k_{\rm B}T \, g_s v \, \lambda_T^{-3} \, b_{5/2}(z(T, v)) & \text{outside the condensation domain} \\ (3/2)k_{\rm B}T \, g_s v \, \lambda_T^{-3} \, \zeta(5/2) & \text{inside the condensation domain} \end{cases} , \tag{381}$$

and in both regions - in the condensation region, because the particles forming the condensate neither exert any pressure nor contribute to energy - the relation $p \, V = (2/3)U$

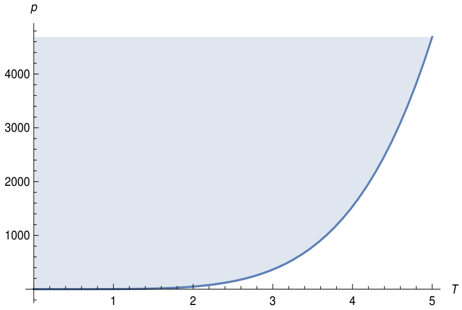

Figure 26: The (T, p) plane phase diagram of noninteracting identical bosons (arbitrary units). The marked region is empty: in the "liquid phase" the pressure is uniquely determined by the temperature, the same as at the "coexistence curve"; therefore the points above this curve do not correspond to any states of the system.

found previously holds true. Owing to this, entropy can be computed as in the case of fermions from the formula

$$\frac{S}{Nk_{\rm B}} = \frac{U - \Omega - N\mu}{Nk_{\rm B}T} = \frac{5}{2}\,\frac{p\,v}{k_{\rm B}T} - \ln z\,,$$

which exploits this relation. This immediately leads to

$$\frac{S}{Nk_{\rm B}} = \begin{cases} (5/2)g_s\,v\,\lambda_T^{-3}\,b_{5/2}(z(T,v)) - \ln z & \text{outside the condensation domain} \\ (5/2)\,g_s\,v\,\lambda_T^{-3}\,\zeta(5/2) & \text{inside the condensation domain} \end{cases}. \tag{382}$$

Entropy can be also computed as

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

Using the first one of the formulae (374) this gives

$$\begin{aligned} S = {} & \frac{5}{2}\,k_{\rm B}\,g_s\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} V\,b_{5/2}(z) - k_{\rm B}\,g_s\ln(1-z) \\ & + \left[k_{\rm B}T\,g_s\left(\frac{mk_{\rm B}T}{2\pi\hbar^2}\right)^{3/2} V\,b'_{5/2}(z) + k_{\rm B}T\,g_s\,\frac{1}{1-z}\right]\left(\frac{\partial z}{\partial T}\right)_{\mu}. \end{aligned}$$

By differentiating with respect to z the series representing the $b_p(z)$ functions in (373) it is easy to find that $b'_{5/2}(z) = z^{-1}b_{3/2}(z)$. Furthermore,

$$\left(\frac{\partial z}{\partial T}\right)_{\mu} = \left(\frac{\partial}{\partial T}\,e^{\mu/k_{\rm B}T}\right)_{\mu} = -\frac{\mu}{k_{\rm B}T^2}\,z = -\frac{z\ln z}{T}\,,$$

so

$$\begin{aligned} \frac{S}{Nk_{\rm B}} = {} & \frac{5}{2}\,g_s\,v\,\lambda_T^{-3}\,b_{5/2}(z) - \frac{g_s}{N}\ln(1-z) \\ & - \left(g_s\,v\,\lambda_T^{-3}b_{3/2}(z) + \frac{g_s}{N}\,\frac{z}{1-z}\right)\ln z\,. \end{aligned}$$

In the thermodynamic limit the term with $\ln(1-z)$ drops out while the entire bracket in the second line equals 1, as follows by comparison with the middle formula in (374). Thus, also this method of calculating entropy leads to (382). As required by 3TMDL, entropy of the system of noninteracting identical bosons vanishes at $T=0$; however, unlike the entropy of the gas of fermions, at very low temperatures entropy of bosons is $\propto T^{3/2}$.

Finally we consider the heat capacity of the system of noninteracting identical bosons. Computation of C_V of the system in the state with the gaseous phase coexisting with the condensate is straightforward because then $z=1$ and entropy is given by the simple formula; therefore:

$$C_V = Nk_{\rm B}\left(T\frac{\partial}{\partial T}\left[\frac{5}{2}\,v\,g_s\,\lambda_T^{-3}\,\zeta(5/2)\right]\right)_{V,N} = \frac{15}{4}\,v\,g_s\,\lambda_T^{-3}\,\zeta(5/2)\,Nk_{\rm B}\,. \tag{383}$$

At very low temperatures the heat capacity vanishes as $T^{3/2}$, in agreement with the requirements imposed by 3TMDL.

Computation of C_V of the pure "gaseous" phase is somewhat more involved (recall: $b'_{5/2}(z)=z^{-1}b_{3/2}(z)$):

$$\begin{aligned}\frac{C_V}{Nk_{\rm B}} &= \left(T\frac{\partial}{\partial T}\left[\frac{5}{2}\,v\,g_s\,\lambda_T^{-3}\,b_{5/2}(z)-\ln z\right]\right)_{V,N}\\ &= \frac{15}{4}\,v\,g_s\,\lambda_T^{-3}\,b_{5/2}(z)+\left[\frac{5}{2}\,v\,g_s\,\lambda_T^{-3}\,b_{3/2}(z)-1\right]\frac{T}{z}\left(\frac{\partial z(T,v)}{\partial T}\right)_v\,.\end{aligned}$$

To work out the second term we rewrite the first bracket in the form

$$\frac{5}{2}\,v\,g_s\,\lambda_T^{-3}\,b_{3/2}(z)-1=\frac{3}{2}-\frac{5}{2}\frac{g_s}{N}\frac{z}{1-z}\,.$$

using the middle one of the three relations (374) and then apply the shocking relation

$$\frac{T}{z}\left(\frac{\partial z(T,v)}{\partial T}\right)_v=\frac{T}{z}\left(\frac{\partial z(T,v)}{\partial T}\right)_{1/v}=-\frac{T}{z}\,\frac{(\partial(1/v)/\partial T)_z}{(\partial(1/v)/\partial z)_T}\,.$$

Also using the middle one of the three relations (374) one can write:

$$\frac{T}{z}\left(\frac{\partial z(T,v)}{\partial T}\right)_v=-\frac{T}{z}\,\frac{(3/2)\,T^{-1}\,g_s\,\lambda_T^{-3}\,b_{3/2}(z)}{g_s\,\lambda_T^{-3}\,b'_{3/2}(z)+(g_s/V)(1-z)^{-2}}\,.$$

Remembering that $b'_{3/2}(z)=z^{-1}b_{1/2}(z)$ and taking into account that in the "gaseous" phase $z\neq 1$, so the terms inversely proportional to N or V vanish in the thermodynamic limit, we finally get

$$\frac{C_V}{Nk_{\rm B}}=\frac{15}{4}\,g_s\,v\,\lambda_T^{-3}\,b_{5/2}(z(T,v))-\frac{9}{4}\,\frac{b_{3/2}(z(T,v))}{b_{1/2}(z(T,v))}\,. \tag{384}$$

Since, as follows from the series in (373), $b_{1/2}(1) = \infty$, the heat capacity per particle is continuous at the "phase transition" point, assuming there its maximal value equal $(15/4)[\zeta(5/2)/\zeta(3/2)]k_{\mathrm{B}} \approx 1.926k_{\mathrm{B}}$ (as can be seen by using the relation (377)), but has there a cusp (the derivatives $(\partial C_V/\partial T)_V$ on both sides of this point are different). In the limit $T \to \infty$, $z \to \lambda_T^3/v \to 0$ (cf. the classical formula (314)), so the expansions $b_p(z) = z + \ldots$ can be used and $C_V/N \to \frac{3}{2}k_{\mathrm{B}}$, in agreement with the classical equipartition theorem.