THERMODYNAMICS AND STATISTICAL PHYSICS
LECTURES

Events important for students of Statistical Physics A (English)

- 2nd December 2019 at 9.00 in room 0.03a - 1st colloquium
- 11th January 2020 at 9.00 in room 1.40 - 2nd colloquium
- 3rd February 2020 at 13.00 in room 1.40 - 1st written exam
- 17th February 2020 at 13.00 in room 1.40 - 2nd written exam

Fizyka Statystyczna A wersja polskojęzyczna

- 25 listopada 2019 o 9.00, sale 1.02 i 2.21 - pierwsze kolokwium
- 13 stycznia 2020 o 9.00, sale B2.38 i 2.21 - drugie kolokwium
- 31 stycznia 2020 o 9.00, sala 0.06 - pierwszy egzamin pisemny
- 20 lutego 2020 o 9.00, sala 0.06 - drugi egzamin pisemny
LITERATURE

- A.B. Pippard, The Elements of Classical Thermodynamics (number 1 of all times)
- 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)
- K. Huang, Statistical Mechanics

and in Polish

- J. Werle, Termodynamika
- K. Gumiński, Termodynamika
- K. Rejmer, Ciepło→Zimno, część pierwsza: Zasady, część druga: Zastosowania
- K. Huang, Mechanika Statystyczna
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 as far as they are influenced by the 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 goals - 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 is important to practically all its areas, from condensed matter to cosmology. Briefly, all physics unites 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 at the Landau Lifschitz theoretical physics course in which the subject is organized in this way) to present this subject mixing statistical physics with thermodynamics. I, however, prefer not to proceed in this way, and would like to clearly separate what is called classical thermodynamics from the statistical physics part in this lecture. Perhaps this will make unhappy some of you - 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 this 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 three statistical ensembles applied mainly to systems of noninteracting particles. Going further would require changing the format of this course to 2x3 at least... But more importantly, a good working knowledge of thermodynamics is important because when the statistical approach is applied to a physical system, the first goal is to recover its thermodynamics and once some basic functions characterizing this system are computed using statistical methods, one applies to them the thermodynamics formalism, so one has first to learn and understand it.

Classical Thermodynamics (CTMD) is a phenomenological theory - it does not enter 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 others 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 specific heats, say \( c_p(T, p_0) \) over a range of temperatures (for a 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 volume will change with temperature, what is its specific heat \( c_v \) for the same range of \( T \) and any volume, compute the work that can be extracted when
the gas is expanded, etc.

CTMD deals only with the equilibrium states of macroscopic physical systems and although considers "processes" in which systems undergo changes, following the time evolution of the systems is beyond its scope. For this reason some would prefer the name thermostatic. Nevertheless, using essentially only idealizations like quasistatic processes, adiathermal isolations, 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 "law" means).

Classical Thermodynamics does not seek to explain the origin of these laws, that is how they emerge as consequences of more fundamental laws which govern the behaviour of microscopic constituents of matter - this is the goal of statistical 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) which again lead to correct predictions in most cases, but here the problem becomes more complicated because application of statistical physics methods relies in most cases on approximations and/or on using simplified models of matter. Nevertheless, analyzing 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 these their general properties which can be treated as macroscopic ones follow only from statistical averaging being to a large extent independent of the details of the microscopic dynamics do indeed obey the laws of TMD.

Of course statistical physics, apart from reproducing thermodynamics and providing inputs which otherwise (within purely thermodynamical treatment) would have to taken from experiments, allows to ask and investigate theoretically questions which are entirely beyond the scope of classical thermodynamics. Here belong for example fluctuations, their spatial correlations, critical exponents (characterizing continuous phase transitions), etc. One can also go to kinetic theories and consider theoretically processes of approaching by systems equilibrium states.

One can think therefore, that classical thermodynamics which was mainly developed when the macroscopic structure of matter was not yet fully investigated and was even questioned is already passé, 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). But
• Frequently, and almost as a rule in applied sciences, it is of primary interest to know relations between properties of substances 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 said: statistical physics will give certain functions - thermodynamical potentials - from which all follows thermodynamically).

• As said, employing statistical physics to real physical systems one is forced to make approximations or use simplified models; knowledge of thermodynamics allows us to tell 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) such systems.

• It sometimes turns out (unexpectedly) that thermodynamics has something to do with the deepest fundamental mysteries - vide entropy of black holes.

• 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 approach 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 culture - for all these reasons it should be an important part of education of a scientist (and any physicist).

“Und somit fangen wir an” - as says Thomass Mann in his Zauberberg.

Two more quotations for the good start:

• 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 begin not to care about. Arnold Sommerfeld

• Every mathematician knows that understanding an elementary course of thermodynamics is impossible. Wladimir Igorevich Arnold

Basic notions
• **Thermodynamic system** - this portion of the Universe which is selected for investigation. A system can be

• **Simple** - a single homogeneous body or a substance; apart from the number(s) (single \( n \) in the case of chemically pure systems - e.g. pure water - and \( n_1, \ldots, n_r \) in the case of a homogeneous mixture of \( r \) components - e.g. a mixture of phenol and water) of moles of its material constituents (some systems, e.g. the electromagnetic field, are not characterized by this variable) only two parameters are needed to fully characterize its equilibrium state - the most widely used example (the “working horse” of thermodynamics) is the gas usually the perfect one) and the variables are then \( p \) and \( V \) (at this stage). One can also consider a paramagnetic substance with the variables \( P \) (total magnetization) and \( H_0 \) or dielectric materials with the variables \( P \) (total polarization) and \( \mathcal{E} \) (when the volume effects can be neglected, e.g. at \( p \approx 0 \), or

• **Not simple** - a single homogeneous body or a substance; more than 2 parameters are needed; it will be seen later that the number of variables (parameters) is equal 1 plus the number of ways a quasistatic work which can be reversibly performed on the system.

• **Non-homogeneous** systems consist of many homogeneous parts (e.g. two containers with same gas or different gases). Sometimes a single nonhomogeneous body (e.g. a gas in the gravitational field) can be mentally split into small parts which can be treated as homogeneous thus allowing to apply to them thermodynamics.

• Under certain conditions a given homogeneous system may split into phases. A phase is a homogeneous part (this means that splitting increases the number of parts of the system) with definite boundaries. A phase may be chemically pure (one \( n \)) or be a mixture - e.g. the phenol-water mixture can split into phases with different concentrations in each.

What is not included in the system constitutes its **surrounding**. Less generally, as the surrounding one takes into account those parts of the Universe which may somehow influence the system under study (here the physical intuition becomes indispensable). In thermodynamical considerations the surrounding is usually taken to consists of sources of work and heat which may be exchanged with the system.

Boundaries separating the system from its surrounding can be natural as in the case of a water droplet or, more frequently, they are artificial or just mental constructions, and are generally called **walls**. They play the role of constraints a 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. For instance, a rigid wall prevents performance of mechanical work on the gas enclosed by it. Walls can also allow or inhibit exchange of matter (of all kind or of only one particular kind) between the system and its surrounding.
Of particular interest for many thermodynamical reasonings are adiathermal walls. Normally we would say such walls inhibit 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 (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). 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 adiathermal in the above sense can allow for some kinds of work to be done on the system, e.g. electrically (by passing a current through a resistant wire inserted in the system), or by stirring.

Thermodynamic variables comprise direct observables (volume $V$, pressure $p$ magnetization $M$ magnetic field strength $\mathcal{H}_0$, etc. and new thermodynamical quantities like temperature $T$, chemical potential(s) $\mu$, enthalpy $H$, entropy $S$... All they fall into two classes intensive ones of essentially local character ($p$, $\mathcal{H}_0$, $\mu$) and extensive ones characterizing the system as a whole (mass $m$, volume $V$, internal energy $U$). If the considered system have the property of extensiveness (most considered systems do have it), it is convenient to operate with specific quantities and molar quantities which are extensive quantities referred to a unit of mass or to one mole, respectively, (sometimes, somewhat incorrectly, molar heat capacities, etc. are also called “specific”). I denote molar quantities by a lower case character, and true specific ones, referred to a unit mass, by a tilde. A quantity which value depends uniquely on the (equilibrium) state of the system is called a state function.

Equilibrium
All thermodynamics rests on the following fact: any system shielded by adiathermal walls and not subject to any external influences tends towards and eventually reaches a state in which no further change (of its macroscopically defined characteristics) is perceptible, no matter how long one waits. This state is called the thermodynamical equilibrium state. It is the state of equilibrium which 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 electric current passing through it. Water flowing down a tube is not strictly speaking in equilibrium in equilibrium - this is the so called steady state - but we will see that to some extent TMD can be applied to it.

If the system is complex (has several parts or some heat or work sources are included
in it), in the approach to equilibria some changes in individual parts occur - they can exchange heat or exchange work between them (e.g. a mercury in glass thermometer inserted in water changes but then it stabilizes) but eventually attain equilibrium. If the external conditions are changed, or walls through which various parts of a complex system communicate, that is the constraints to which the system is subjected have changed, a new equilibrium state is attained and one of the main roles of thermodynamics is to determine this state (the Callen’s point of view).

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

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

- unstable mechanical - no analogy.

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

- metastable mechanical - supercooled vapour or a mixture of hydrogen $H_2$ and oxygen $O_2$: 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 they are in fact not in equilibrium states. For many practical purposes can be, if not artificially stimulated, treated as being in full equilibrium.

In fact no unstable equilibrium exists in mechanics too: it is a purely mathematical concept, for it relies on many idealizations (pointlike material points, neglect of internal structure of bodies, neglect of minute external influences). Real physical systems can have only metastable equilibria but the range of displacements which do not make them changing the 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 the system realized in specific conditions (specific constraints) minimize or maximize a suitable functions called thermodynamics potentials (examples are entropy, Helmholtz free energy, Gibbs potential, etc).

As a matter of facts, equilibrium in thermodynamics is never static: microscopic examination of particles or molecules in the system always reveals they are in a state of continuous agitation (Brownian motion); local density of a fluid fluctuates a little around the mean value If one waits long enough - theoretically, because the times involved are longer than the Universes lifetime! - one might observe 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 physics (or even the kinetic theory of gases). And from it, in addition to being able to compute such numbers, one draws a lesson that such fluctuations should be treated as immanent part of the equilibrium state. Anyway, for most purposes after reasonably short time one can treat any system as having attained equilibrium state. If the consequences of treating it as such are not corroborated by experiment, we must revise this assumption - evidently the system has not attained equilibrium or its fluctuations play a crucial role in the phenomenon we want to explain.

0TMDL and Temperature

To introduce and discuss this concept, fundamental to 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 - deformation of the shape without any change in volume do not require any work. The shape of a solid can only be changed by the application of stress and its thermal properties are then usually also affected.

We accept a fact of experience that the equilibrium state of a fixed mass of a fluid is completely specified by its volume $V$ and pressure $p$ (we assume it has no electric or magnetic properties or that elemg fields are absent altogether).

We can fix the volume $V$ of the fluid kept in a cylinder and adjust $p$ to any desired value by placing the fluid 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. Anything macroscopic that can be measured as the characteristic of this system can be treated as a unique function of $V$ and $p$.

Take now any two simple systems (two fluids for definiteness). If they are isolated and reach equilibrium states separately, and then are brought into thermal contact through a diathermal (rigid, so that mechanical work is not possible) wall, changes in general will be observed to occur in both until they reach, as a compound system, a new equilibrium state. So we say they are now in thermal equilibrium with one another.

On the basis of our sensory experience we say that this is so because initially they in general do not have same temperature. But this word has no content yet. What is 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. They in general will not be in equilibrium with one another unless $V_1$, $p_1$, $V_2$ and $p_2$ are 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 (no changes in their states are observed when they are brought into thermal contact). So, their equilibrium requires that a relation

$$ F(V_1, p_1, V_2, p_2) = 0, \quad (1) $$

be satisfied. 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 show
that this relation always takes the form

\[ \phi_1(V_1, p_1) = \phi_2(V_2, p_2), \]  

with the function \( \phi_1 \) being the property of only the first system and \( \phi_2 \) of only the second
one. This is shown relying on

**0TMDL**

*If of the three bodies A, B and C, A and B and are separately in thermal equilibrium with
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 diather-
mal walls and the whole system in equilibrium then any two of them taken separately are
also in equilibrium with one another.*

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

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

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
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 the 0TMDL the same isotherm of the
system \( T \) will be obtained.

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\( t \), so that we define a function (which need not be analytic at this stage)

\[ \phi_T(V_T, p_T) = t, \]

and we call \( t \) the **empirical temperature**. Once this is done, we can produce isotherms
of all other 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
body must be associated a function \( \phi_{\text{body}} \) depending on the parameters fully characterizing
equilibrium states of that body and this function must take values equal \( \phi_T(V_T, p_T) \) 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 - (empirical) temperature. From this it also clear that
if more than two parameters are needed to completely characterize equilibrium states of

10
a body, say \(x_1, \ldots, x_r\), its isotherms are not curves, but hypersurfaces of dimension \(r - 1\) (or of codimension 1) determined by the equation

\[
f_{\text{body}}(x_1, \ldots, x_r) = t. \tag{3}
\]

Now the mathematical reasoning. Consider three fluids \(A\), \(B\) and \(C\). If \(A\) and \(C\) are in equilibrium, then

\[
F_1(V_A, p_A, V_C, p_C) = 0, \quad \text{so} \quad p_C = f_1(V_C, V_A, p_A).
\]

Similarly, if \(B\) and \(C\) are,

\[
F_2(V_B, p_B, V_C, p_C) = 0, \quad \text{so} \quad p_C = f_2(V_C, V_B, p_B).
\]

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 \(B\) also are in equilibrium so

\[
F_3(V_A, p_A, V_B, p_B) = 0,
\]

and (4) must be equivalent to this. But \(V_C\) does not enter \(F_3\), so it must drop out from \(f_1 = f_2\). 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 \(\psi\) and \(\eta\).

That 0TMDL enforces dropping out of \(V_C\) 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 \(f_1 = f_2\) 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 we can fix it and forget it changing the notation to

\[
f_3(V_B, V_A, p_A) \equiv g_A(V_A, p_A).
\]

so 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

\[
p_C = f_2(V_C, g_A(V_A, p_A)),
\]

and 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 obtains two other relations

\[
h_A(V_A, p_A) = h_B(V_B, p_B), \quad \text{and} \quad t_B(V_B, p_B) = t_C(V_C, p_C),
\]
but we don’t know yet, e.g. how the function \( t_B(V_B, p_B) \) relates to \( h_B(V_B, p_B) \). One has to show that these three relations between 6 functions can be reduced to three relations involving only 3 functions each depending on only one pair of variables \( V, p \).

To this end from \( g_A(V_A, p_A) = g_C(V_C, p_C) \) we get \( p_C = \psi(V_C, V_A, p_A) \) and put it into

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

We now take \( V_A, V_B, V_C \) and \( p_A \) for independent variables - in other words we treat \( p_B \) as determined by these four) and differentiate both sides of the above equality w.r.t. \( V_B \). Because the right hand side is independent of \( V_B \), this gives

\[
\frac{\partial t_B}{\partial V_B} + \frac{\partial t_B}{\partial p_B} \frac{\partial p_B}{\partial V_B} = 0 .
\]

In turn, differentiating w.r.t. \( V_B \) the relation between the functions \( h_A \) and \( h_B \) (continuing to treat \( V_A, V_B, V_C \) and \( p_A \) as independent variables) we get

\[
\frac{\partial h_B}{\partial V_B} + \frac{\partial h_B}{\partial p_B} \frac{\partial p_B}{\partial V_B} = 0 .
\]

Eliminating now the derivative \( \partial p_B/\partial V_B \) we find

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

Vanishing of this Jacobian identically means in effect that the mapping \( (V_B, p_B) \rightarrow (t_B, h_B) \in \mathbb{R}^2 \) is degenerate (is of rank 1, instead 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 this curve and this relation can be again 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) or because one can always use \( \dot{t}_A = f(t_A) \), \( \dot{t}_B = f(t_B) \) and \( \dot{t}_C = f(t_C) \) (in the mathematical reasoning) the quantity \( t \) (a state function) is called empirical. We will see that TMD itself provides a mean of defining the **absolute** temperature \( T \) with the scale
being the only freedom in its definition. The question therefore is how to relate empirical
temperatures defined by different thermometric bodies to the absolute 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 $x$ as $x = f(t)$ 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 between them or one ascribes a concrete
value $t$ to one particular point.
For instance, one takes the Mercury in glass at pressure $p = 0$, sets $h_{Hg} = a_{Hg} t_{Hg}$ ($h_{Hg}$
being the height - above some conveniently chosen reference level - of the column of
Mercury in evacuated glass tube at $p = 0$) and determines $a_{Hg}$ by requiring that there
be 100 degrees between the melting point of ice and the boiling point of water, both at
normal pressure $1.013 \times 10^5$ Pa (the famous 1013 HPa $\equiv 1$ atm). However, if one takes
another thermometric substance, e.g. the ethyl alcohol and constructs $h = a \tilde{t}$ in the same
way, both thermometers can be made (by the appropriate choices of the reference levels of
the height $h_{Hg}$ and $h$) to give $t_{Hg} = \tilde{t}$ at the two chosen reference points but will in general
differ (even if not too much in practice) at all other point ($t_{Hg} \neq \tilde{t}$). This is because these
are two different empirical temperatures.

Another choice of the thermometric substance is gases at very low pressures. This is
simple because the isotherms of gases under this condition are simple: $pV = \text{const}$. This
is the Boyle-Marriott law which is satisfied to a very good accuracy by real gases at
sufficiently low pressures. So one can set $pV = nRt$ and determine $nR$ as above. It then
turns out that $t$ and $\tilde{t}$ obtained with different gases (at sufficiently low pressures) are
nearly the same not only at the reference points but in a wide range of temperatures.
(Moreover, it happens experimentally that if the empirical temperature $\tilde{t}$ is fixed by the
Mercury in glass thermometer by the rule $\tilde{t} = a h$, then over a wide range of temperatures:
$\tilde{t} \approx t$. ) This is of course because the temperature defined in this way using the perfect
gas (a theoretical construct) is exactly proportional to the absolute temperature $T$ (the
one determined by the 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 it is
easier to reproduce in laboratories the water triple point the definition was changed and
now the absolute temperature scale is fixed by ascribing to this point (which corresponds
to $p = 611.73$ Pa) $T = 273.16$ 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 absolute
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.

Once the absolute temperature is established, the commonly used Celsius scale is defined as $t = T - 273.15$. On this scale the triple point corresponds to $0.01^\circ C$. 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 by $h = at + b$ and ascribing $0^\circ C$ to the ice melting point (and therefore $100^\circ C$ to the water boiling point). 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 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 $(13.81 \text{ K})$, triple point of Oxygen $(54.361 \text{ K})$, the melting point of Zinc at 1 atm $(692.73)$ melting point of Gold $(1337.58 \text{ K})$.

Thermometers based on expansion of liquids (Mercury, Ethyl Alcohol, Pentane) cover the range of $-300^\circ C$ up to $+300^\circ C$.

Resistant thermometers are based on the variation of electrical resistance of a metal with temperature cover even larger range. E.g. Platinum is easy to purify, has high melting point $(1770^\circ C)$; it is very accurate between $70 \text{ K}$ and $1500 \text{ K}$.

Thermocouple thermometers use variation of the e.m.f. with $T$. One keeps one junction at fixed temperature and the e.m.f depends on the temperature of the other one. Small voltages have to be measured - this make them difficult to work with if high accuracy is needed but such thermometers can be miniaturized and respond quickly to changes of $T$.

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

Carbon resistors. Useful below $20 \text{ K}$. Below $10 \text{ K}$ also have sensitivity of order $10^{-5} \text{ K}$.

Below $5 \text{ K}$ to somewhat below $1 \text{ K}$ liquid $^4\text{He}$ is used. Between $1 \text{ K}$ and $0.3 \text{ K}$ $^3\text{He}$ is used. In both cases $T$ is found by measuring the vapour pressure.

For even lower $T$ one relies on paramagnetic salts. Their susceptibility goes like $a/T$ (Curie law)

Above the Gold melting point only measurements of radiation emitted by hot bodies are used (radiation pyrometers).

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.

Finally it should be stressed that the mere introducing 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 mean yet 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 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 in its state is perceptible macroscopically and on OTMDL (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 obvious mechanical ones like pressure $p$ or volume $V$) state function - 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 when they are contacted with one another through a rigid (preventing their mechanical contact) diathermal wall impermeable for matter. Such two systems are then said to be in thermal equilibrium with one another. (Notice, that $t_A = t_B$ is not sufficient for full thermodynamical equilibrium which 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 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 $t$: it can be always replaced by $\tilde{t} = f(t)$ where $f(\cdot)$ is any monotonic function. But once this arbitrariness of $t$ has been fixed (by adopting some its definition - however arbitrary - by choosing a thermometric body), it is a single valued function of the parameters $x_1, x_2, \ldots$, like $p, V$, and other ones, like the amount of matter in the system represented by the number of moles $n_1, \ldots, n_r$ of its $r$ components needed to fully characterize the equilibrium state:

$$t = t(x_1, \ldots, x_s, n_1, \ldots, n_r). \quad (5)$$

Such a functional relation is called the equation of state of the system (of the body, of the substance). Naturally, thermodynamics by itself does not determine its form and it has to be determined experimentally or else derived using the statistical physics approach.

We recall also 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).

**Internal energy**

We begin by introducing the notion of internal energy. Let us consider the historic experiments performed by Joule which originally were intended to measure the mechanical equivalent of heat (but for us the word “heat” has as yet no content, so we will interpret this experiment somewhat differently).
One realization of this sort of experiment can be the original Joule's paddle-wheel immersed in an (as ideally as possible) adiathermally isolated calorimeter containing a mass of a liquid (e.g. water). A measurable mechanical work can be performed on this system by rotating the wheel by an angle \( \alpha \) 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 \( g \) propelling the wheel: measuring its final speed \( w \) one gets the work \( mgh - mw^2/2 \) done on the system (the liquid) by the device. The state of the liquid is found to change (its temperature changes, its pressure changes) as a result of performing the work.

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

Similar experiments, employing different kind of directly measurable work can be performed. The important fact is that 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 (I would add here “equilibrium states” - see the footnote below).

This is the 1TMDL as applied to adiathermally isolated systems.

Example: the state of a gas adiathermally isolated (against any influences from without) changes from \( A \) to \( B \). Path 1: an amount of electric work is first performed isochorically on it, until it reaches the state \( Y \); then the gas is adiathermally 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 and it attains the same final state \( B \). Then the statement is that

\[
W_{A \rightarrow Y \rightarrow B} = W_{A \rightarrow X \rightarrow B}.
\]

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 on 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.\(^1\)

\(^1\)We require the states \( A \) and \( B \) to be equilibrium states because only then they can be specified by giving the values of a few parameters only; the statement remains true also if one or both these states are not equilibrium states but it would then be harder to be sure that the same state \( B \) has been reached on both paths.
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, owing to 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 the 1TMDL applied to adiathermally isolated systems one can introduce 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 performance on it an amount $W$ of work, its internal energy shall be said to have increased by

$$\Delta U \equiv U_B - U_A = W \quad \text{on adiathermal paths}. \quad (6)$$

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

$$U_A = U_0 + W_{R \to A} \quad \text{on adiathermal paths}. \quad (7)$$

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 $A$ to another state $B$ or the other way around, that is from $B$ to $A$. 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 $\Delta U$ it is sufficient that it can be traced in one way only. Since I always had trouble to understand what this statement means, I advice you to think again on the paddle-wheel device: call the state $A$ the state of the fluid when the mass $m$ is in the higher position and $B$ the state when the mass $m$ is in the lower one. The path $A$ to $B$ can be realized (in fact in numerous ways). The path $B$ to $A$ evidently not and not only with the help of the paddle-wheel device but by any other means, so long as the liquid is adiathermally isolated.

In any case, the important thing is that differences of internal energies in various states and therefore, this quantity up to an additive constant, can be measured by measuring only mechanical or electrical works (not heat) and that $U$ is a function of state, so it can be expressed (in case of equilibrium states) as a function of the parameters needed to specify this state:

$$U = U(p, V), \quad \text{or} \quad U = U(p, t) = U(p, V(t, p)), \quad \text{or} \quad U = U(t, V) = U(p(t, V), V),$$

in the case of a simple fluid, etc. In this way, on the manifold of equilibrium states of the system (parametrized by some convenient state variables) we superimpose the network of internal energies. Of course, microscopically this internal energy (by
definition considered in the body’s rest frame) consists of kinetic and interaction energies of the matter constituents. (tu może o addytywności U)

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

**Heat as the “work defect” and TMDL**

Once a value of internal energy $U$ is unambiguously (when the reference state $R$ is chosen) assigned to any (equilibrium) state in the way sketched above, one can consider changes of the system during which it is not adiathermally isolated. It is then possible to accomplish a change of the system’s state from $A$ to $B$ in more different ways and some of these ways involve amounts of work which are different than $\Delta U = U_B - U_A$. For instance, in the considered example of a simple fluid taken from $A$ 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 work. So in this case $\Delta U = U_B - U_A \neq W_{A\rightarrow Y\rightarrow B}$. We then define the quantity $Q$ by

$$Q = \Delta U - W_{A\rightarrow Y\rightarrow B},$$

and call it (somewhat misleadingly, but this should cause no harm - we are already far from the discussions of XIX century!) heat taken by the fluid 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 great contribution to teaching thermodynamics! - **Heat is the “defect” of work**

by obvious analogy to the binding energy defined as the defect of mass in nuclear physics (nuclei weight less than protons and neutrons out of which they are made and this deficit, called 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). 1) when absorbed 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), 2) may be conveyed from one body to another by conduction, convection and radiation (obvious - even vacuum does not entirely inhibit changes effected by other means than work - changes caused by radiation can be inhibited by walls perfectly reflecting electromagnetic waves) 3) 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 and one is asked to compute the final temperature, or given some data, the heat capacities etc.)

Only the property 3 requires some brief justification: consider a typical calorimetric experiment in which two bodies, 1 and 2, at different temperatures $t_1 \neq t_2$ are brought into a thermal contact within a calorimeter (a Dewar vessel) whose walls are adiathermal and rigid (no work can be performed in the bodies within the calorimeter by external agents). Therefore the total change $\Delta U$ of the system’s internal energy must be zero

$$\Delta U = \Delta U_1 + \Delta U_2 = 0.$$
But
\[ \Delta U_1 = W_1 + Q_1, \quad \Delta U_2 = W_2 + Q_2, \]
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
\[ Q_1 + Q_2 = 0. \]

In calorimetric experiments (but not in general!) heat is conserved.
Thus we finally arrive at the **general 1TMDL** in the form\(^2\)
\[ \Delta U = W + Q. \]

In words: energy is conserved if heat \( Q \) is included and recognized as a form of energy transferred from one system to another. Callen (whose celebrated textbook on thermodynamics will be exploited later) describes the difference between energy transferred through mechanical work and in the form of heat as the results of couplings of different types of degrees of freedom: 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 heat is due to the coupling of microscopic degrees of freedom of the system and its environment (this pictorial explanation does not, however, seem to apply to work performed in magnetizing or polarizing electromagnetically active bodies).

The equivalence of heat and work can be illustrated by the example of a gas enclosed in a cylinder with piston. The system is here the gas and the cylinder and the whole system is adiathermally isolated. If the gas expands doing some work \( W \) on the piston (not necessarily reversibly - we still do not know what this word means) and next the whole work \( W \) 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 (its state will not in general be the initial one).

**Notions “hotter”, “colder” and their correlation with the scale of \( t \)**
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 losing. This is called **heat transfer** though this does not imply existence of heat as a physical entity (“caloric” or “flogiston”) whose movement can be followed from one body to another.

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

\(^2\)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 of these quantities can be either positive or negative.
fact, following from experience, that the rate of the heat transfer may usually be varied
over a wide range, depending on the nature of diathermal wall through which the bodies
exchange heat (the rate is a measure of thermal conductance of the wall).

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

So we now will 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 $t = t(p, V, \ldots)$ 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
made precise above) than all bodies at temperature $t_1$ if $t_2 > t_1$.

This is done 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 $A$.

It is then possible the break somewhat 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 $t_A \neq t_B = t_C$; note we do not say
$t_A$ is greater or lower than $t_B = t_C$).

If the three bodies are then brought to 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 it is possible to define the empirical tempera-
ture so that if $t_2 > t_1$ than any body at $t_2$ is hotter than any body at $t_1$.

In the reasoning we have assumed (analyzing the heat conservation in a calorimeter ex-
periment) we have allowed for the possibility that the two bodies do work on one another
(depending on other properties of the wall through which they are in contact, while they
exchanged heat between them). Therefore, if those parameters which together with $t$ are
used to specify the states 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 them
is replaced by a rigid movable one) The correlation of $t$ and hotness will still obtain. The
consequence of this is that absorption of heat by a body whose independent parameters
other than $t$ remain constant will always cause an increase of $t$. Therefore the so called
principal heat capacities (like $C_V$, $C_p$) characterizing the 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 (conven-
tional) assignement 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 during 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 which conditions in an infinitesimal change of the system’s state from one equilibrium state to another one the 1TMDL

\[
dU = q + w, \tag{13}
\]

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 calls them in thermodynamics) heat and work,\(^3\) can be written with \( w \) and \( q \) being forms (differential one-forms in thermodynamics called Pfaffian forms) on the space of the parameters \( x_1, \ldots, x_s \) characterizing equilibrium states 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 \( df_Q(x_1, \ldots, x_s) \) of \( df_W(x_1, \ldots, x_s) \), that is, as exact differentials (or closed forms) because in finite changes the heat and work 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\(^4\)

\[
\hat{\omega}_W \equiv dW, \tag{14}
\]

so that in these conditions also \( q \) can be written as

\[
\hat{\omega}_Q \equiv dQ. \tag{15}
\]

Thus, if the necessary conditions are met, one will have the right to write

\[
dU = dQ + dW. \tag{16}
\]

(But, as we are going to discuss in the next Lecture, owing to 2TMDL the character of the form \( dQ \) is different than that of \( dW \).)

As usually it is convenient to consider a simple system, a gas enclosed in a cylinder fitted with a movable piston of cross section area \( A \). In the equilibrium state the forces acting on the piston must be balanced so that the net force is zero. These forces are: a)

---
\(^3\)We do not consider here the possibility that separately \( w \) and \( q \) are large but mutually cancel out leaving only an infinitesimal sum \( w + q \).

\(^4\)Thermodynamical tradition requires the differential forms which are not exact differentials of state functions be written with the slashed \( d, \dd \) (Kubo uses \( d'W \)), instead of \( \hat{\omega} \) (which is the mathematical notation). Because when I started to write notes for classes and problems I did not know how to produce this symbol in LaTeX, in some places \( dW \) or \( dQ \) may still be found (instead of \( d'W \) and \( dQ \)). While I will do my best to find these places and replace the symbols, I hope that students who encounter \( dW \) or \( dQ \) will be smart enough to not take them for exact differentials. (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...)}
the force which the gas exerts on the piston, equal \( pA \), where \( p \) is the gas pressure, b) an external force \( F_{\text{ext}} \) applied to the piston from outside (can be provided by the pressure of the environment or by other means). When the piston starts to move, however slowly, also frictional force can enter into play. Now suppose, the piston has moved by \( dx \) while being acted upon by \( F_{\text{ext}} \). The work done by the external force on the system (the gas, the cylinder and the piston) certainly is

\[
F_{\text{ext}} \cdot dx = -F_{\text{ext}} dx = -\frac{F_{\text{ext}}}{A} d(Ax) \equiv -p_{\text{ext}} dV .
\] (17)

The work done on the system, assuming the absence of friction, is this minus the kinetic energy acquired by the piston on the distance \( dx \).

In general \( p_{\text{ext}} \) 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 it by two parameters only: the gas pressure will not be the same in the entire gas volume - it will be different from point to point. Only if the piston moves very slowly, without acquiring any acceleration, so that at any moment the gas inside the cylinder remains practically in equilibrium and one can ascribe to it a unique pressure \( p \) (almost the same in the whole volume) and there is no friction, can one say that \( p_{\text{ext}} = p \) and write the work done by the external force on the distance \( dx \) in the form

\[
dW = -pdV , \quad \text{or} \quad dW = -p(t,V) dV , \quad \text{or} \quad dW = -p(V,U) dV ,
\] (18)

depending on the choice of independent variables - \((p,V)\), or \((t,V)\) or \((U,V)\) - characterizing the equilibrium states. If this is possible, the work \( q \) 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 \( dW = -pdV \) is the limiting case of a process which can be practically realized.

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

- the change must occur vanishingly slowly - the process must be \textbf{quasistatic}, meaning that at each stage the system can be treated as if it were in full equilibrium (within itself and with its surrounding)

- there should be no friction - the process must be \textbf{reversible} which in practical terms means that it occurs under an \textit{infinitesimal} difference between \( p \) and \( p_{\text{ext}} \); infinitesimal change of this difference (which reverses its sign) suffices to change the \textbf{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 limiting 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. 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. 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 process will be \textbf{quasistatic}. The product \( -pdV \) will certainly be non-zero. Yet it will not represent the work done on or by the gas - the work will be exactly zero! One can also make in this way a finite change of the volume occupied by the gas, just by opening to it successively and always quasistatically, additional volumes \( dV \) until a finite \( \Delta V \) is reached. The work done by the gas will still be zero irrespectively of the fact that \( \sum(-pdV) = -\int dV p \neq 0 \). 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.\footnote{As always, there is a lot of confusion in terminology here. Many authors, including my respected colleagues from the 5th floor, 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.}

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 gas is maintained at each stage. When the gas is being compressed (the volume \( V \) decreases), \( p_{\text{ext}} \) 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, \( p_{\text{ext}} \) 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 (make a plot)

\[
W = -\int dV p_{\text{ext}} > 0.
\]  

If the final gas temperature at the end of this process is the same as the initial state (we assume that the volume after the cycle returns to its initial value), this means that \( W \) 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) \( W \) is just the difference of the gas final and initial energies. In both situations we have to do with 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 the work done by this force is not given by

\[
-\int dV p,
\]  

so \( w \) - the work done on the gas cannot be written as \( -pdV \).
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 infinitesimal change of the temperature of one of these bodies would reverse the direction of the heat flow. This is the always the case when isothermal processes are considered (provided other elements of such processes are carried out reversibly, there is no friction, etc.): one assumes that the temperature $t$ of the system, while it undergoes such a change during which heat is added to or extracted from it, remains equal to the temperature $t_{\text{res}}$ of a reservoir (called also heat bath) which, owing to its very large size (infinite in the limit), stays unchanged (its temperature does not change and its equilibrium is not perturbed), now matter how much 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 during 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.

Thus when the fluid undergoes a reversible change, one can write
\[ dU = q - p \, dV. \] \hspace{1cm} (21)

This means that in such a change also $q$ can be written as $dQ$, 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 other way as $dU + p \, dV$. This 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 ground for problems which will be assigned as class, home and colloquial works.

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

  \[ dW = \gamma \, dA, \] \hspace{1cm} (22)

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

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

  \[ dW = K \, dL. \] \hspace{1cm} (23)

  $L$ is here the rubber length and $K$ its strain. The relation $K = K(T, L)$ plays the role of the equation of state.
• Work needed to stretch a thin wire.

\[ dW = K dL, \]

where \( L \) is the wire length and \( L \) its tension. This can be reversible if the deformation is elastic (and not plastic; 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 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 distances in the body after the deformation are encoded in the tensor

\[ u^{ij} = \frac{1}{2} \left( \frac{\partial u^i}{\partial x^j} + \frac{\partial u^j}{\partial x^i} + \sum_l \frac{\partial u^l}{\partial x^i} \frac{\partial u^l}{\partial x^j} \right) \approx \frac{1}{2} \left( \frac{\partial u^i}{\partial x^j} + \frac{\partial u^j}{\partial x^i} \right), \]

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

\[ dV' = dV \left( 1 + \text{tr}(u^{ij}) \right) = dV \left( 1 + \lambda^{(1)}(\mathbf{x}) + \lambda^{(2)}(\mathbf{x}) + \lambda^{(3)}(\mathbf{x}) \right). \]

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

\[ dW = \int_V d^3\mathbf{x} \sigma^{ij}du^{ij}, \]

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

\[ u^{ij} = \frac{1}{9K} \delta^{ij} \sigma^{ll} + \frac{1}{2\mu} \left( \sigma^{ij} - \frac{1}{3} \sigma^{ll} \sigma^{ij} \right) \]

or

\[ \sigma^{ij} = K u^{ll} \delta^{ij} + 2\mu \left( u^{ij} - \frac{1}{3} u^{ll} \delta^{ij} \right). \]

where \( K \) is the compressibility modulus and \( \mu \) 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 \( \sigma^{ij} = -p \delta^{ij} \) and the work reduces to \( -pdV \). (See L&L, vol. 8, Elastomechanics.)

• Work needed for magnetization. Magnetic materials (paramagnetic or diamagnetic) put in an external magnetic field of strength \( \mathcal{H}_0 \) (the subscript 0 indicates this is the
magnetic field produced by the currents (in a coil, for example) acquire local magnetization $\mathbf{M}(x)$ 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 homogeneous in space magnetic field strength)

$$dW = \mathbf{H}_0 \cdot d\mathbf{M},$$

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

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

$$dW = \mathbf{E} \cdot d\mathbf{P},$$

where $\mathbf{P} = \int d^3x \, \mathbf{P}$. Again the relation between $\mathbf{P}$ and $\mathbf{E}$ plays the role of the equation of state.

A digression

Before we move further it is useful to make a short summary. We started with characterizing equilibrium states of any system by parameters directly related to works which can be performed on the system reversibly. In the case of a fluid these are $V$ and $p$, or $\mathbf{M}$ and/or $\mathbf{P}$ if systems with nontrivial electromagnetic properties are considered. 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) 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 $\mathbb{R}^n$. A coordinate system is therefore a set of $n$ independent functions (their number $n$ 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 has been chosen, other functions like $U$, $t$ or $V$ become functions of the coordinates used to identify the 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 - the same can happen - si puo dare (compressed
to “podarsi”) as Gioelle Botta 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 $dW$, $dQ$ 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 the **Enthalpy** $H = U + pV$, the **Helmholtz free energy** $F = U - TS$ or the **Gibbs function** $G = U - TS + pV$. They are called **thermodynamic potentials** ($U$ and entropy $S$ - to be introduced in the next Lecture - also are called thermodynamic potentials) for in specific situations they indeed play analogous role the potentials $V(q^1, \ldots, q^s)$ play in Mechanics - they determine the equilibrium state of the system. As state functions, all of them can be expressed in terms of different state parameters.\(^6\)

Let us briefly demonstrate the usefulness of enthalpy in chemistry. Most of chemical reactions occur at constant (atmospheric) pressure. During such 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 $p_{\text{ext}}$). If the initial and final states of the system are in equilibrium within themselves and with the surrounding atmosphere ($p = p_{\text{ext}}$, $t = t_{\text{ext}}$), 1TMDL can be written as $U_{\text{fin}} = U_{\text{in}} + W + Q$, 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 $-p_{\text{ext}}(V_{\text{fin}} - V_{\text{in}})$, and in the initial and final states $p = p_{\text{ext}}$, one can write

$$U_{\text{fin}} = U_{\text{in}} - p (V_{\text{fin}} - V_{\text{in}}) + Q,$$  \hspace{1cm} (32)

or

$$ (U_{\text{in}} + pV_{\text{in}}) - (U_{\text{fin}} + pV_{\text{fin}}) \equiv H_{\text{in}} - H_{\text{fin}} = Q. $$  \hspace{1cm} (33)

Thus the heat $Q$ released in a 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 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 in various flow processes - its conservation generalizes to compressible fluids the Bernoulli law (which applies only to incompressible fluids).

\(^6\)Later we will see, however, that each of these thermodynamic potentials has a preferred (in the sense that will be elucidated) set of its variables.
LECTURE III (TMD)

We have already introduced and discussed two of the four laws of thermodynamics. In the preceding Lecture it was found 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 \( dw \) (or \( dw \), 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 \( dw \) are well defined differential forms, also \( q \) must be a differential form \( dQ \). However, on the basis of what has been done, there is no way to write this differential form differently than as \( dQ = dU - dw \), for example, considering a simple fluid, as

\[
dQ = dU(p, V) + p\, dV,
\]

if the variables \( p \) and \( V \) are taken for independent ones. 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

\[
Q = \int_A^B dQ = \int_A^B (dU(p, V) + p\, dV),
\]

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

Adiabatic changes of the system, that is, (in my terminology) its reversible adiabatic changes, are those changes in which the form \( dQ \) vanishes. More mathematically, \( dQ \) 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 \((p, V)\) or \((p, t)\) or \((t, V)\) spaces determined by the solutions of the differential equation

\[
dU + p\, dV = 0,
\]

written in the set of variables, \((p, V)\) or \((p, t)\) or \((t, V)\), which has been chosen (on the basis of convenience) to work with, and an initial point (state). In the general case of a system whose equilibrium states are determined by the parameters \( x_1, \ldots, x_s \) and \( X_i \) are the generalized forces which can act on the system, adiabats are curves (paths) on which projections of the one-form\(^7\)

\[
dQ = dU(x_1, \ldots, x_s) - \sum_{i=1}^{s-1} X_i(x_1, \ldots, x_s)\, dx_i,
\]

vanish.

\(^7\)Recall, that the number \( s \) of independent parameters \( x_i \) equals 1 plus the number of works which can be reversibly done on the system.

29
Unfortunately in practice one usually does not know $U$ as the function of the system’s parameters like $p$ and $V$, or $x_1, \ldots, x_s$ 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.\footnote{However, some of the thermodynamical relations \textbf{can} be obtained with this limited knowledge which is encoded in 0k1TMDLs - see the home and colloquia problems to this course.}

At this point it is amusing to say that the perfect gas (and 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 its internal energy $U$ is independent of its volume\footnote{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 positions of molecules), changes of the volume do not alter $U$.} $V$, the main obstacle for playing with this system is removed, because passing to $t$ and $V$ as the independent variables, the heat form of the perfect gas can be explicitly written down as

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

(38)

$C_V^{(t)}$ is here the heat capacity absorbed by the system when its empirical temperature $t$ changes by one unit, but since usually one includes in this definition 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 inconsciously) replaces $t$ by $T$ has no consequences.

The adopted definition of the perfect gas allows to solve problems of the sort “what heat it 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 in the isothermal expansion and, by 1TMDL, \(Q = -W = \bar{W}\), where

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

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 (38), on isotherms, owing to the assumption that $(\partial U/\partial V)_t = 0$, the heat form of the perfect gas takes the form $dQ = p \, dV$. 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 such a gas does depend on the volume $V$.

As a matter of facts, as a consequence of 2TMDL, 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 its internal energy $U$ on its volume $V$ is inconsistent with the perfect gas equation of state $pV \propto t \propto T$ (fortunately, it is consistent, as you will be able to check).

Furthermore, in reality, independence of temperature of the heat capacity $C_V$ (2TMDL does not constrain its dependence on the temperature) is true only if the perfect gas is a monoatomic one - $C_V = n c_v$ in this case equals $\frac{3}{2} n R$, where $n$ is the number of moles. Gases, 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 temperature is lowered) the three-dimensional motion of gas molecules 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 motion (rotations and vibrations) of compound molecules, $C_V/n$ of multiatomic gases rises to $3R$ ($\frac{3}{2} R$ 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 ($+1 R$ 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. 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

$$T \sim E_{exc}/k_B,$$

where $E_{exc}$ are typical energies of the rotational and vibrational excitations of the gas molecule and $k_B$ is the Boltzmann constant, $^{10}$ $k_B = 8.617 \times 10^{-5}$ eV/K. Typical energies of rotational excitations of molecules lie in the $3 \times 10^{-5} \div 5 \times 10^{-3}$ eV while those of vibrational excitations in the range $0.1 \div 1$ eV (energies of the rotational excitations are lower than energies of the vibrational ones - 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 tens to hundreds Kelvins (excitations of rotations) and then in the range of thousands Kelvins; in the intermediate temperatures there can also be some departures from the constancy of specific heats due to e.g. fine structure of the atomic ground states (the corresponding energies are of order $10^{-2}$ eV). Thus, at least some quantum effects manifests themselves in the specific heats of gases only at very high (compared to the room ones) temperatures (not only, as one naively could think, at very low temperatures)!

$^{10}$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...
2TMDL

Only some of 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 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).\footnote{It is perhaps fair to say already in this place, that within the statistical approach it turns out that at the microscopic level 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 - in statistical approach the discussed tendency of changes to occur in only one direction is a statistical effect. But probabilities of such returns are so fantastically tiny (one talks here of numbers as small as $10^{-10^{n-tens}}$) that they never happen in practice and are never observed. Thus, phenomenological thermodynamics which is based on what is really (and not what hypothetically could be) observed can safely rest on the discussed assumption.}

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 $dW$ 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 $p = p_{\text{ext}}$, $t = t_{\text{ext}}$, quasistaticity etc.). Normal and prevailing type of behaviour of real systems are 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 2TMDL 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):

\textit{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 \textit{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.

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 takes 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, 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.

\textbf{2TMDL (Kelvin, 1851, again in Pippard’s words):}

\textit{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.}

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:

\textit{It is impossible to absorb heat from a reservoir and to convert it all into work without introducing any other changes in the system and its environment.}

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 proved to be equivalent. This is a sort of a scholastic exercise, which I recall quickly here.

1. Clausius—\( \Rightarrow \) Kelvin. We argue that if the Kelvin’s principle were violated, one could violate also the Clausius one (proof by \textit{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 \(^{12}\) which takes heat \( \bar{Q}_1 \) from the reservoir at \( t_1 \) and transfers a positive amount \( Q_2 = \bar{Q}_1 + \bar{W} \) of heat to the

\(^{12}\) It is assumed here that everybody knows what the Carnot cycle or engine is. (Carnot cycles will appears here three pages further). 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 doing work!
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 $Q_2 > \bar{Q}_1$ from the reservoir at $t_2$, give back the heat $\bar{Q}_1$ to the reservoir at $t_1$ producing the net work $\bar{W} = Q_2 - \bar{Q}_1$, in effect entirely at the cost of heat taken from the reservoir at $t_2$.

There is one more formulation of 2TMDL, due to Caratheodory, 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, Caratheodory 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. Caratheodory, 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 Caratheodory’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\(^{13}\) and, therefore, states of lower internal energies are inaccessible; here however, the Caratheodory’s principle is assumed to apply to any thermodynamical system, no matter how complicated.

It is fairly straightforward to show that the Caratheodory’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 Caratheodory’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 the system make an isothermal change after which it reaches a state $B$ in this domain, absorbing on the way a positive heat $Q$ (from a reservoir).\(^{14}\) The assumption that the Caratheodory’s principle is violated in the considered domain means, that all points are accessible from $B$ on adiathermal paths. So $A$ must be accessible too. If the system has returned to $A$ by an adiathermal path, the total change of its internal energy is zero and, by 1TMDL, the positive heat $Q$ which it absorbed from the reservoir in the isothermal transition from $A$ to $B$ must have been 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 $dQ$ of a system is identically zero. Then any two states within this domain can be connected by an adiathermal path, but vanishing of $dQ$ means that the system is entirely mechanical

---

\(^{13}\) Adiathermal expansion of the liquid could decrease (a bit) its energy, but in this way the liquid would not return to the initial state which is, therefore, inaccessible adiathermally from the final state.

\(^{14}\) 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.
and not thermodynamical, so in this domain 1TMDL reduces $\Delta U = W$. Such systems indeed do not satisfy the Caratheodory’s principle but this does not invalidate, of course, 2TMDL in the Kelvin’s and Clausius’ formulations.)

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 experience. We will come to discuss this formulation a bit later.

The mathematical consequences of 2TMDL are the following

- The differential one-form of heat, $dQ$, of any thermally homogeneous system, no matter how complicated,\(^{15}\) 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 $dQ$ there is one, $1/T$, which is given by a universal (i.e. independent of the system) function $T(t)$ of the empirical temperature $t$.

- The exact differential $dS = dQ(x_1,\ldots,x_r)/T(t(x_1,\ldots,x_r))$ defines, up to a constant, a new state function of the system - its entropy $S(x_1,\ldots,x_s)$.

- In all adiathermal processes (reversible or not) the entropy does not decrease: $S_{\text{fin}} \geq S_{\text{in}}$.

- 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 Caratheodory’s formulation of 2TMLD seems to be rather abstract but it allows to derive its mathematical consequences without appealing to cyclic processes whose use (outside the theory of machines producing work or refrigerators) is rather artificial and for this reason 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 the 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 colleagues from the 5th floor) for the Caratheodory’s principle directly postulates (through the second Caratheodory’s theorem) the existence of adiabatic surfaces, that is, it directly states

\(^{15}\)As explained below, this statement is trivial in the case of simple systems whose equilibrium states are characterized by two parameters only.
that the heat forms $dQ$ of physical systems do have integrating factor and are therefore integrable, with values of the entropy $S$ labeling families of the corresponding solutions of the equations $dQ = 0$. But, at least to me, the whole essence of thermodynamics is to see 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 $S_{\text{fin}} \geq S_{\text{ini}}$ determines the direction of real processes occurring in isolated 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.

- 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 many other places.

To quote R. Emden: “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.”
LECTURE IV (TMD)

We now set ourselves to infer the integrability of \( dQ \) and, in consequence, the existence of the entropy as a state function, 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).

At the beginning we will consider adiabatic (adiathermal reversible) changes in which \( dQ = dU - dW = 0 \). In the case of simple systems, e.g. fluids, characterized by the variables \( p \) and \( V \), the equation

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

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

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

determines uniquely in the parameter space a family of nonintersecting curves\(^{16}\) labeled by the initial point which therefore makes the heat form \( dQ \) trivially integrable in such cases. If the system is the perfect gas, the equation reduces to

\[
C_V^{(t)} dt + p(t, V) dV = 0,
\]

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 \( V^{nR/C_V^{(t)}} = \text{const.} \), 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.) The heat form \( dQ \) is made integrable just by dividing it by \( t \) (or const. \( \times t \)), 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).

However in classes you have seen (at least those who attended...) an example (taken from Pippard’s book) of a one-form in three dimensions which is not integrable. Thus not every one-form in more than two dimensions is integrable.

To establish integrability of the form \( dQ \) in the general case, we will (as in the case of proving the existence of the empirical temperature as a function of state) first use a reasoning which is more physical (and uses virtually no mathematics) and then a more mathematical one.

To show physically, using the Carnot cycles, that the Kelvin’s 2TMDL implies integrability of \( dQ \), we start from two isothermal hyper-surfaces \( t(x_1, \ldots, x_s) = t_1 \) and

\(^{16}\)It should be known from Math II course (at least to those trained in math classes by me) that the equations of the type \( F(x, y)dx + Q(x, y)dy = 0 \) sometimes do have singular points through which more than one integral curve (solution) passes; such cases are here excluded.
\( t(x_1, \ldots, x_n) = t_2 \) with \( t_1 \neq t_2 \) (if the system has three parameters these are ordinary two-dimensional surfaces in the three-dimensional parameter space). From 0TMDL we know such hyper-surfaces always exist and are not intersecting (the empirical temperature is a unique state function - it does not have “branches”). We consider on the surface \( t = t_1 \) a curve connecting two points \( A \) and \( B \) belonging to this surface. Then we consider two curves which are two solutions of the equation \( dQ = 0 \) passing through the points \( A \) and \( B \) and intersecting the surface \( t = t_2 \) at the points \( D \) and \( C \), respectively (The equation \( dQ = 0 \) can always be integrated step by step starting from the points \( A \) and \( B \) and continuing until it crosses the hypersurface \( t = t_2 \), although the continuation may not be unique). In this way one obtains a \textbf{reversible cycle} - the Carnot cycle, called also the Carnot engine - whose segments \( AB \) and \( CD \) correspond to isothermal changes of the considered system while its segments \( BC \) and \( DA \) correspond to adiathermal reversible changes. Beginning from the state \( A \), the system takes in the isothermal change \( A \to B \) a heat \( Q_1 \) 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 \( t_2 \). 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

\[
\bar{W} = Q_1 + Q_2 .
\]  

(43)

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 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 with the help of two different thermodynamical systems working between the same two empirical temperatures \( t_1 \) and \( t_2 \), of which the first engine absorbs heats \( Q_1 \) and \( Q_2 \) and the other one \( Q'_1 \) and \( Q'_2 \). 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 \( k'/k \) (Cauchy sequences!). One can then treat the two Carnot engines as a single 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 taken from the reservoir at \( t_1 \) is 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

\[
k Q_2 + k' Q'_2 \leq 0 .
\]  

(44)
But because the individual cycles, and therefore the cycle of the compound system, are reversible, also
\[-(kQ_1 + k'Q'_2) \leq 0,
\]
which follows from the possibility of accomplishing the cycle of the compound system in the opposite sense. Thus, \(kQ_1 + k'Q'_2 = 0\) and from the two equalities
\[kQ_1 = -k'Q'_1,
\]
\[kQ_2 = -k'Q'_2,
\]
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 \(t_2\):
\[-Q_1/Q_2 = f(t_1, t_2),
\]
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 consisting of one Carnot cycle \(C_{12}\) working between \(t_1\) and \(t_2\), consisting as previously of the changes \(A \to B \to C \to D\), and another one, \(C_{23}\), working between \(t_2\) and \(t_3\) and consisting of the changes\(^{17}\) \(D \to C \to F \to G\). Viewing the compound cycle composed of the cycle \(C_{12}\) followed by the cycle \(C_{23}\) as a single cycle executed between the temperatures \(t_1\) and \(t_3\), on one hand one has \(-Q_1/Q_3 \equiv -Q_{1(2)}^{(12)}/Q_{3}^{(23)} = f(t_1, t_3)\) while on the other hand
\[-Q_1/Q_3 = f(t_1, t_3) = \left(\frac{-Q_{1}^{(12)}}{-Q_{2}^{(12)}}\right) \left(\frac{-Q_{2}^{(23)}}{-Q_{3}^{(23)}}\right) = f(t_1, t_2) f(t_2, t_3),
\]
because \(Q_{1}^{(12)} = -Q_{2}^{(23)}\). 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,
\[-Q_1/Q_2 = \phi(t_1)/\phi(t_2) \equiv T_1/T_2.
\]
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 is universally given by the ratio of the temperatures defined in this way of the two reservoirs, \(T\) is called the absolute temperature. Thus\(^{18}\)
\[T = \text{const.} \times \phi(t),
\]
\(^{17}\)It is not necessary that these two cycles be performed by the same system; it is enough that the heats \(Q_{1}^{(12)}\) 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.
\(^{18}\)As discussed, the proportionality constant in (49) 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 1: A realization of the fragment $\Gamma_i$ of the cycle executed by the body $\sigma$.

and operationally the temperature $T$ is determined by the efficiency

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

of any Carnot cycle working between $t_1$ and $t_2$. This means that any empirical temperature $t$ defined using a thermometric body can 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 to be proportional to the absolute one, it is easier (if the physical conditions allow for this) to relate a given absolute temperature to the perfect gas temperature.

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

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0.$$  \hfill (51)

We now generalize this result, still relying on the Kelvin’s formulation of 2TMDL.

Consider a system $\sigma$ 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) $\Gamma_i$, $i = 1, \ldots, M$ ($M \to \infty$). At the $i$-th subprocess $\Gamma_i$ of the cycle a work is done on or by the system $\sigma$ and some heat is transferred to or abstracted from this system. In the gedanken experiment this transfer of the amount $q_i$ 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 reservoir at a temperature $T_i$ (see Figure 1). One can imagine that the Carnot cycle $C_i$ delivers the amount $q_i$ 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 $\sigma$. The last step - the transfer of $q_i$ to $\sigma$ may be irreversible; we do not assume that the system $\sigma$ 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 (51), from $R_0$ the heat

$$q_i \frac{T_0}{T_i}.$$  

(52)

Moreover, since all Carnot cycles $C_i$ return to their initial states, and the same is assumed about the system $\sigma$, the total amount $\dot{W}$ 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 Kelvin’s principle, the work $\dot{W}$ cannot be positive (for this would simply mean taking an amount of heat from $R_0$ and converting it all into work). Thus

$$\dot{W} = \sum_i q_i \frac{T_0}{T_i} \leq 0.$$  

(53)

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

$$\oint \frac{q}{T_{\text{ext}}} \leq 0,$$

(54)

known as the Clausius inequality\(^{19}\) (despite the fact that we derived it by relying on the Kelvin’s principle).\(^{20}\) The temperature under the integral has been here appended the subscript “ext” to stress (strongly!) the fact that this is not the temperature of the system $\sigma$ (this may not be definable) but the temperature of the (changed in the process) reservoirs from which the heat $q$ is supplied to the system $\sigma$ and which all are in equilibrium within themselves (they can be assumed to be sufficiently large).

Only if the cycle executed by the system $\sigma$ is reversible, can one identify the temperatures $T_{\text{ext}}$ with the actual (on a given stage of the system’s $\sigma$ cycle) temperature of the system $\sigma$ (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 $\sigma$ is reversible, it can be executed in the opposite sense, leading to the inequality

$$-\oint \frac{q}{T_{\text{ext}}} = -\oint \frac{q}{T} \leq 0.$$  

(55)

In this case, because the changes of the system $\sigma$ composing the cycle it executes are reversible, the elementary heat $q$ can be interpreted as the differential form $dQ$ on the space of the system’s $\sigma$ equilibrium parameters (and, as said, $T_{\text{ext}}$ identified with the actual system’s temperature $T$). Hence, if the cycle is reversible, the Clausius inequality becomes the equality

$$\oint \frac{dQ}{T} = 0,$$

(56)

\(^{19}\)This inequality is frequently written erroneously with the sign $\ge$, because students usually remember that changes of entropy are nonnegative and the factor under the integral $q/T_{\text{ext}}$ is confused with $dS = dQ/T_{\text{ext}}$.

\(^{20}\)In classes you will prove it relying directly on the Clausius’s own principle.
which generalizes the equality \( Q_1/T_1 + Q_2/T_2 = 0 \) holding for Carnot cycles.

The last equality allows to define entropy \( S \) as a state function, because it shows that the integral of \( dQ/T \) taken between the equilibrium states \( A \) and \( B \) of any thermodynamical system does not depend on the reversible path between these states. If \( \Gamma_{A\rightarrow B}^1 \) and \( \Gamma_{A\rightarrow B}^2 \) are two such paths, then

\[
0 = \int \frac{dQ}{T} = \int_{\Gamma_{A\rightarrow B}^1} \frac{dQ}{T} + \int_{\Gamma_{A\rightarrow B}^2} \frac{dQ}{T},
\]

hence

\[
\int_{\Gamma_{A\rightarrow B}^1} \frac{dQ}{T} = \int_{\Gamma_{A\rightarrow B}^2} \frac{dQ}{T}.
\]

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

\[
S_B - S_A = \int_A^B \frac{dQ}{T},
\]

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, then the entropy of any other state \( A \) can be defined as

\[
S_A = S_0 + \int_R^B \frac{dQ}{T},
\]

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 (it is reasonable to take the same reference state \( R \) for \( U \) and \( S \) by linking the states with adiathermal paths.

The existence of entropy as a function of state means that the heat form \( dQ \) is integrable - the surfaces which are solutions of the equation (recall the material of classes!) \( dQ = 0 \) are simply the surfaces \( S = \text{const.} \) and its integrating factor, which is singled out by its dependence on the systems parameters \( x_1, \ldots, x_s \) only through the empirical temperature \( t(x_1, \ldots, x_s) \) is just the absolute temperature \( T = \phi(t(x_1, \ldots, x_s)) \). In reversible adiathermal processes entropy \( S \) stays constant. This also means that in a reversible processes the heat form \( dQ \) can be written as

\[
dQ = TdS.
\]

Using this relation which must hold, it is possible to show that if the integration factor of the form \( dQ \) is restricted to depend only on \( t \), the only its nonuniqueness and nonuniqueness of entropy reduces to

\[
\bar{T} = aT, \quad S = \frac{1}{a} S + b.
\]
Indeed if alternative temperature $\tilde{T}$ and entropy $\tilde{S}$ are introduced so that $dQ$ can be written in two ways as

$$dQ = T \, dS = \tilde{T}(T) \, d\tilde{S}(x_1, \cdots, x_s),$$

(since both, $T$ and $\tilde{T}$ must be functions of $t$ only, $\tilde{T} = \tilde{T}(T)$), one can pass to the new set of variables (coordinates on the manifold of states) with $S$ being one of them and $y_1, \ldots, y_{s-1}$ the remaining ones. In these new variables the above equality can be written as

$$dS = \frac{\tilde{T}(T)}{T} \left[ \left( \frac{\partial \tilde{S}}{\partial S} \right)_{y_i} \, dS + \sum_{i=1}^{s-1} \left( \frac{\partial \tilde{S}}{\partial y_i} \right)_{S,y_{i+1}} \, dy_i \right].$$

From this, by comparing the coefficients of the differentials on both sides, it readily follows that $(\partial \tilde{S}/\partial y_i)_{S,y_i} = 0$, which implies that $\tilde{S} = \bar{S}(S)$ and, moreover, that

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

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 $T/\tilde{T}(T) = 1/a = d\bar{S}/dS$. This gives the result.

We can now present another way of arriving at the same consequences of 2TMDL (still relying on the Kelvin’s principle) without using the 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 $dQ = Y_1 dx_1 + Y_2 dx_2 + Y_3 dx_3$. We begin by drawing in the system’s parameter space an adiabat $\Gamma$ (constructed by making steps correlated by $dQ = 0$) crossing “successive” (they are distributed continuously) isothermal surfaces. Since the system has three parameters, it is possible to draw lines beginning at the adiabat $\Gamma$ 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 $\Gamma$. The integral

$$\int_P^{P'} dQ,$$

taken along the constructed adiabatic lines (along the two ones lying on the isothermal surfaces and the adiabat $\Gamma$) 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 \( dQ = 0 \) (more precisely, the one-form \( dQ \) 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 \( dQ \) 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 surfaces. Moreover, since now we know that on the vector \( PP' \) tangent to the surface the form \( dQ \) gives zero, it 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 (if it could, then this path closed with \( PP'' \) would be a reversible cycle violating the Kelvin’s principle). So the surface is entirely surrounded by points inaccessible on reversible adiabathermal paths. Other adiabatic surfaces can be constructed similarly starting from any other curve along which \( dQ = 0 \). The reasoning can be generalized to more complicated systems requiring more than three parameters (one constructs in such cases adiabatic hypersurfaces of dimension \( s - 1 \).

The existence of adiathermal surfaces can be also demonstrated taking as the starting point the Caratheodory’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 Caratheodory’s principle states that in the neighbourhood of any point (state) of any thermodynamical system there are points 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 the general 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 representing a reversible process connecting \( P \) with \( Q \), there would have to exist a point \( P' \), arbitrarily close to \( Q \), accessible on adiabatic 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 infinitesimally close to it (point \( P' \)) but not infinitesimally close to \( P \) - some points would be then distinguished. Hence, the Caratheodory’s principle requires in fact that infinitesimally close to any state of any thermodynamical system there are points inaccessible on adiabatic (reversible adiathermal) paths. The construction of adiabatic surfaces proceed then almost as previously: starting from any point \( P \) one draws a \( dQ = 0 \) line through it, then an adiabatic surface, containing this line, then the three-dimensional hypersurface containing this surface, and so on. Any surface can be then cordoned from both sides by other not intersecting it such hypersurfaces constructed by starting from points infinitesimally close to \( P \) on both sides of the constructed hypersurface and inaccessible from it adiabatically.

As the adiabatic surfaces can be constructed (the approaches based on Kelvin’s and
Caratheodory’s principles merge at this point), one can now proceed to showing that in a reversible change the heat form $dQ$ can be written as $TdS$ with $T$ being the function of the empirical temperature $t$ only. One starts by labeling the constructed adiabatic (hyper)surfaces of dimension $s-1$ by a parameter $\sigma$, constructing thereby a function $\sigma(x_1, \ldots, x_s)$, much in the same way as the function $t(x_1, \ldots, x_s)$ has been introduced by labeling different isothermal surfaces (and in direct analogy with $t$ the function $\sigma(x_1, \ldots, x_s)$ can be called the empirical entropy).

As the adiabatic surfaces exist, from the first Caratheodory’s theorem\(^{21}\) it follows that the form $dQ(x_1, \ldots, x_s)$ 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 $dQ = 0$, the function $\sigma(x_1, \ldots, x_s)$ is constant, in any adiabatic (reversible adiathermal) change of the considered system

$$d\sigma \equiv \sum_{i=1}^{s} \frac{\partial \sigma}{\partial x_i} dx_i = 0,$$

(62)

similarly as\(^{22}\)

$$dQ = \sum_{i=1}^{s} X_i \, dx_i = 0.$$

(63)

Therefore,

$$dQ - \lambda d\sigma \equiv \sum_{i=1}^{s} \left( X_i - \lambda \frac{\partial \sigma}{\partial x_i} \right) \, dx_i = 0,$$

(64)

where $\lambda(x_1, \ldots, x_s)$ can be an arbitrary state function (on the space of states of the system). In this relation only $s-1$ differentials $dx_i$ are independent (can be changed at will), because they are (at every point) correlated by (62). If one regards, $dx_1$ as determined by the remaining differentials, one can adjust the arbitrary state function $\lambda(x_1, \ldots, x_s)$ so as to make the coefficient of $dx_1$ in (64) vanish\(^{23}\)

$$X_1(x_1, \ldots, x_s) = \lambda(x_1, \ldots, x_s) \frac{\partial \sigma}{\partial x_1}.$$

(65)

Then, since (64) must hold for any adiathermal reversible change in which $dx_2, \ldots, dx_s$ can be chosen arbitrarily, one concludes that

$$X_i(x_1, \ldots, x_s) = \lambda(x_1, \ldots, x_s) \frac{\partial \sigma}{\partial x_i},$$

(66)

\(^{21}\)Hopefully discussed in classes.

\(^{22}\)Here $X_i$ are not necessarily the generalized forces because we included in them also the terms following from $dU(x_1, \ldots, x_s)$.

\(^{23}\)And this is the essence of the trick with the Lagrange multipliers.
now for all \( i = 1, \ldots, s \). Hence

\[
dQ = \sum_{i=1}^{s} X_i(x_1, \ldots, x_s) \, dx_i = \sum_{i=1}^{s} \lambda(x_1, \ldots, x_s) \frac{\partial \sigma}{\partial x_i} \, dx_i = \lambda \, d\sigma .
\] (67)

1/\( \lambda \) is therefore an integrating factor of \( dQ \). One has, however, to show (and this does not follow directly from the Carathéodory’s first theorem) that \( \lambda \) cannot be an arbitrary function of the parameters \( x_1, \ldots, x_s \) but necessarily takes a form of a product of two functions: one which depends on \( x_1, \ldots, x_s \) through the empirical temperature, and another one which depends on these parameter through the function \( \sigma \) itself.

To this end we consider two systems, 1 and 2, the first one characterized by the variables \( x_1, \ldots, x_s \) and the second one by \( y_1, \ldots, y_r \), each of which has its own heat form (\( dQ_1 \) and \( dQ_2 \)), its own system of labeling adiathermal (hyper)surfaces (functions \( \sigma_1(x_1, \ldots, x_s) \) and \( \sigma_2(y_1, \ldots, y_r) \)) and an integration factor of its heat form (functions \( \lambda_1(x_1, \ldots, x_s) \) 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

\[
t_1(x_1, \ldots, x_s) = t = t_2(y_1, \ldots, y_r) .
\] (68)

The compound system, being thermally homogeneous, is therefore characterized by \( s + r - 1 \) variables for which one can take the common temperature \( t \) and \( x_2, \ldots, x_s \) and \( y_2, \ldots, y_r \). As any thermally homogeneous thermodynamical system it too must have its function \( \Sigma(t, x_2, \ldots, x_s, y_2, \ldots, y_r) \) and its heat form \( dQ \) must have an integrating factor \( \Lambda(t, x_2, \ldots, x_s, y_2, \ldots, y_r) \). Since \( dQ = dQ_1 + dQ_2 \), the following relation must, therefore, hold

\[
\Lambda(t, x_2, \ldots, x_s, y_2, \ldots, y_r) \, d\Sigma(t, x_2, \ldots, x_s, y_2, \ldots, y_r) = \lambda_1(t, x_2, \ldots, x_s) \, d\sigma_1(t, x_2, \ldots, x_s) + \lambda_2(t, dy_2, \ldots, y_r) \, d\sigma_2(t, y_2, \ldots, y_r) .
\]

Dividing both sides by \( \Lambda \) and going over to the variables \( t, \sigma_1, \sigma_2, x_3, \ldots, y_3, \ldots \) one has

\[
d\Sigma(t, \sigma_1, \sigma_2, x_3, \ldots, y_3, \ldots) = \frac{\lambda_1}{\Lambda} \, d\sigma_1 + \frac{\lambda_2}{\Lambda} \, d\sigma_2 .
\]

It is now clear that \( \Sigma \) is a function of \( \sigma_1 \) and \( \sigma_2 \) 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 \( \lambda_1 \) does not depend on \( y_3, \ldots, y_r \), neither can \( \Lambda \), and in the same way, since \( \lambda_2 \) does not depend on \( x_3, \ldots, x_s \), neither can \( \Lambda \). It then follows, that

\[
\lambda_1 = \lambda_1(t, \sigma_1) , \quad \lambda_2 = \lambda_2(t, \sigma_2) , \quad \Lambda = \Lambda(t, \sigma_1, \sigma_2) ,
\]

and, moreover, the dependence on the empirical temperature \( t \) must also drop out from the ratios \( \lambda_1/\Lambda \) and \( \lambda_2/\Lambda \):

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

46
and from this it follows that
\[
\frac{\partial}{\partial t} \ln \lambda_1(t, \sigma_1) = \frac{\partial}{\partial t} \ln \lambda_2(t, \sigma_2) = \frac{\partial}{\partial t} \ln \Lambda(t, \sigma_1, \sigma_2).
\] (69)

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
\[
\frac{\partial}{\partial t} \ln \lambda_1(t, \sigma_1) = g(t), \quad \frac{\partial}{\partial t} \ln \lambda_2(t, \sigma_2) = g(t)
\]
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
\[
dQ_i = \lambda_i \, d\sigma_i = \left[ a \exp \left( \int dt \, g(t) \right) \right] \frac{1}{a} W_i(\sigma_i) \, d\sigma_i,
\] (70)

It remains to show that the integrating factor \(\Lambda\) of the compound system also has this structure, with the same universal function of the empirical temperature and the function \(W\) which depends on \(\sigma_1\) and \(\sigma_2\) only through \(\Sigma\). From (69), 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 \(dQ = dQ_1 + dQ_2\), one can write
\[
W(\sigma_1, \sigma_2) \, d\Sigma(\sigma_1, \sigma_2) = W_1(\sigma_1) \, d\sigma(\sigma_1) + W_2(\sigma_2) \, d\sigma(\sigma_2),
\]
from which it is clear that
\[
W(\sigma_1, \sigma_2) \frac{\partial \Sigma}{\partial \sigma_1} = W_1(\sigma_1), \quad W(\sigma_1, \sigma_2) \frac{\partial \Sigma}{\partial \sigma_2} = W_2(\sigma_2).
\]

It is now sufficient to differentiate the first of these two equalities with respect to \(\sigma_2\) and the second one with respect to \(\sigma_1\) to arrive at the relation
\[
\frac{\partial W}{\partial \sigma_1} \frac{\partial \Sigma}{\partial \sigma_2} - \frac{\partial W}{\partial \sigma_2} \frac{\partial \Sigma}{\partial \sigma_1} = \frac{\partial (W, \Sigma)}{\partial (\sigma_1, \sigma_2)} = 0,
\]
which means that the mapping of \((\sigma_1, \sigma_2)\) into \((W, \Sigma)\) 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,
\[
W(\sigma_1, \sigma_2) \, d\Sigma = W(\Sigma(\sigma_1, \sigma_2)) \, d\Sigma,
\]

47
which is just what we wanted to prove.

Thus the heat form \( dQ \) of any thermally homogeneous thermodynamical system can be written as

\[
dQ = \lambda \, d\sigma = \left[ a \exp\left( \int dt \, g(t) \right) \right] \frac{1}{a} W(\sigma) \, d\sigma \equiv T(t) \, dS,
\]

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. \( S \) therefore the true entropy of the system.

Summarizing, in reversible changes the heat form of any thermodynamic system can be written as \( dQ = TdS \) and in such changes

\[
dU = TdS + \sum_{i=1}^{s-1} X_i \, dx_i. \tag{71}
\]

This differential relation can be now extended to all differential changes, reversible ones and also irreversible ones, because it is simply the relation between the state functions at infinitesimally close points (representing equilibrium states of the system). However one should remember, that only in reversible changes the differentials \( TdS \) and \( X_i \, dx_i \) 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

\[
dQ = TdS, \quad dW = \sum_{i=1}^{s-1} X_i \, dx_i. \tag{72}
\]

To give an example: consider a gas expanding into an additional volume of the cylinder when the gas (the whole cylinder) is adiathermally isolated. As discussed, this change can be accomplished quasistatically, successively opening additional volumes \( dV \), 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 opened additional 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 cylinder 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 is energy \( U \) or anything else) in a process \( A \to B \) in which the system passes from an equilibrium state \( A \) to another equilibrium state \( B \), but which process by 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.

Having introduced entropy as a state function, we can express the heat capacities of systems in reversible processes in a convenient way. Recall that in any quasistatic change (in which the system passes through a sequence of equilibrium states) the system’s heat capacity $C$ is defined by projecting the form $dQ$ onto the curve in the space of states representing this process:

$$dQ|_{\text{projected}} = C(x_1(T), \ldots, x_s(T)) \, dT.$$  

(73)

If the change is reversible (and only then) one can replace $dQ$ by $TdS$ and write

$$C(x_1(T), \ldots, x_s(T)) = T \frac{d}{dT} S(x_1(T), \ldots, x_s(T)).$$

(74)

In most cases the process is specified by constancy of certain parameters, e.g. $V$ or $p$ in the case of simple fluids, and this is written as

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

(75)

One more important thing follows from the presented reasonings: in the case of systems in thermal equilibrium (like the two systems considered above), that is having the same $t$ and therefore $T$,

$$dQ = dQ_1 + dQ_2 = TdS_1 + TdS_2 = Td(S_1 + S_2) = TdS,$$

(76)

so the entropy is additive. If the systems are not at the same temperatures but are not in thermal contact (are separated by an adiathermal wall) one defines the entropy of the compound system as $S = S_1 + S_2$. 

49
LECTURE V (TMD)

Accepting 2TMDL as a generally valid law of Physics which applies to any macroscopic thermodynamical system, we have shown that the heat form $dQ$ of any (thermally homogeneous) system can, in reversible processes, be written as $TdS$, where $T = T(t(x_1, \ldots, x_s))$ is the absolute temperature (so, the indicator telling us whether two system brought into a thermal contact will be in equilibrium) and $S(x_1, \ldots, x_s)$ the new state function - the entropy. In infinitesimal reversible processes 1TMDL, $dU = q + w$, can therefore be written as the sum

$$dU = TdS + \sum_{i=1}^{s-1} X_idx_i,$$  \hspace{1cm} (77)

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. We have in this way obtained a convenient representation of heat capacities characterizing the system in reversible processes as derivatives of its entropy

$$C_X = T \left( \frac{\partial S}{\partial T} \right)_X.$$  \hspace{1cm} (78)

We have also stressed that the formula (77) 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 (77), which in this case takes the form

$$dU = TdS - p\,dV,$$  \hspace{1cm} (79)

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 (79) that

$$T = \left( \frac{\partial U}{\partial S} \right)_V, \quad p = -\left( \frac{\partial U}{\partial V} \right)_S.$$  \hspace{1cm} (80)

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

$$\left( \frac{\partial T}{\partial V} \right)_S = -\left( \frac{\partial p}{\partial S} \right)_V.$$  \hspace{1cm} (81)

This is a relation 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 heat which is measured; in this way the change $\Delta S$ of the fluid’s entropy can be determined as $\Delta S = \Delta Q/T$; if this is divided by the measured resulting change
\( \Delta p \) of the fluid's pressure, the right hand side is (although not so easily, as is always the case with calorimetric measurements) obtained. The coefficient on the left hand side is much easier to measure: one simply measures the change \( \Delta T \) of the fluid's temperature resulting from a change \( \Delta V \) of the fluid's volume in the adiabatic (strictly speaking adiabatic) conditions.

The relation (81) 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 one 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 \( dQ \); from the physical point of view they are highly nontrivial consequences of 2TMDL whose validity 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 \( P \) are needed and the fundamental relation (77) reads \( (P \) is the system's total polarization vector and \( \mathbf{E} \) is the electric field in which the system is placed)

\[
dU = TdS - pdV + \mathbf{E} \cdot d\mathbf{P}.
\]  

Instead of considering the internal energy \( U \) of this system, one can consider another thermodynamic potential (a state function), call it \( \Phi \), defined as

\[
\Phi = U - ST + pV - \mathbf{E} \cdot \mathbf{P}.
\]

It is straightforward to see that

\[
d\Phi = -SdT + Vdp - \mathbf{P} \cdot d\mathbf{E}.
\]  

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

\[
\left( \frac{\partial \Phi}{\partial T} \right)_{p,\mathbf{E}} = -S, \quad \left( \frac{\partial \Phi}{\partial p} \right)_{T,\mathbf{E}} = V, \quad \left( \frac{\partial \Phi}{\partial \mathbf{E}} \right)_{T,p} = -\mathbf{P}.
\]

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

\[
\left( \frac{\partial \mathbf{P}}{\partial p} \right)_{T,\mathbf{E}} = -\left( \frac{\partial V}{\partial \mathbf{E}} \right)_{T,p}.
\]

This shows that there must be a nontrivial relation between two seemingly unrelated phenomena: piezoelectricity - generation of a polarization of a material as a result of squeezing it - and electrostriction - which is the change of the material's volume in the electric field. 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. Two identical homogeneous balls $A$ and $B$ of mass $M$ made of, say, iron or another material of nonnegligible thermal expansivity property, have the same initial temperature. One of them, $A$ rests on a horizontal plane (in the Earth’s gravitational field $g$), say, on a table, and the other one, $B$, is suspended on a thread (see the left Figure 2). The same quantities $\delta Q$ of heat are supplied to both balls. Which one will have higher temperature? This problem was once assigned at the International Physics Olympiad and then appeared in many sources. The expected answer was that 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 its 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 $\delta Q$ is used up for 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 $\alpha^{\text{lin}} = (1/R)(dR/dT)$ be the linear expansion coefficient of the ball and $C_0$ its heat capacity (which we can take to be independent on the temperature; for simplicity we can assume that all the experiment is carried out at zero pressure) in the absence of the gravitational field. Then

$$
\text{ball } A : \quad \delta Q = (C_0 + MgR\alpha^{\text{lin}}) \delta T_A , \\
\text{ball } B : \quad \delta Q = (C_0 - MgR\alpha^{\text{lin}}) \delta T_B ,
$$

from which it readily follows 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 2): 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 (directed downwards) and negative (upwards) values. Using this convention one can introduce $C(g) \equiv C_0 + MgR\alpha^{\text{lin}}$ as the ball’s heat capacity in the gravitational field. Then

$$
\left( \frac{\partial C}{\partial g} \right)_T = MR\alpha^{\text{lin}} .
$$

---

24 $\Phi$ is the Legendre transform of $U$. We will tell more about this transformation later.

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 3. Start with the ball on the table at a temperature $T_1$. Then bring it (not changing its position in the gravitational field) into the thermal contact with a heat bath of temperature $T_2 = T_1 + \delta T > T_1$. According to the discussed “solution”, 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 lose 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 $T_2 > T_1$ goes into increasing the balls potential energy by $2\delta R$. This potential energy can be converted into a mechanical work $\delta \dot{W}$ 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 $\eta$ of the cycle? This is easily computed as the ratio of the work done to the heat absorbed:

$$\eta = \frac{\delta \dot{W}}{\delta Q_{\text{abs}}} = \frac{2MgR\alpha_{\text{lin}}\delta T}{(C_0 + MgR\alpha_{\text{lin}})\delta T} = \frac{2MgR\alpha_{\text{lin}}}{(C_0 + MgR\alpha_{\text{lin}})}.$$ 

It is clearly independent of the temperature difference $\delta T$ of the two reservoirs! But according to 2TMDL the efficiency of any cycle 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 $\delta T$ goes to zero. Choosing $\delta T$ sufficiently small, on could, if the presented “solution” were right, beat the efficiency of the Carnot cycle!

What was wrong? We have assumed that the ball in the gravitational field does not get deformed - that its shape is 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 independent 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 of state

\[
f(T, Y, g) = 0.
\]

Applied to the ball 1TMDL takes the form

\[
d\delta Q = \delta U + Mg\delta Y,
\]

or, if reversible changes are considered, the familiar form

\[
dU = TdS - Mg\delta Y.
\]

(The sign of the second term is the same as in 1TMDL, \(dU = TdS - p\,dV\) applied to a fluid because in increasing \(Y\) the ball must do a positive work against gravity - if \(g > 0\) - 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 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 \(H = U + MgY\), whose differential is

\[
dH = TdS + MY\,dg = T \left( \frac{\partial S}{\partial T} \right)_g dT + \left[ T \left( \frac{\partial S}{\partial g} \right)_T + MY \right] dg
\]

\[
= \left( \frac{\partial H}{\partial T} \right)_g dT + \left( \frac{\partial H}{\partial g} \right)_T dg.
\]

Its second form shows that \(H\) is treated here 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:

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

\[
\left( \frac{\partial C_g}{\partial T} \right)_T = -MT \left( \frac{\partial^2 Y}{\partial T^2} \right)_g \equiv -MTY \left[ \alpha^2_{\text{lin}} + \left( \frac{\partial \alpha_{\text{lin}}}{\partial T} \right)_g \right],
\]

(87)

This is markedly different than the naive formula (86)! In particular, the sign of the derivative is opposite if, as usually happens with real materials, \( (\partial \alpha_{\text{lin}}/\partial T)_g > 0 \) (or at least if this derivative is not too large negative to outweigh the positive \( \alpha^2_{\text{lin}} \) term). Therefore, the heat capacity \( C \) of the ball is larger if \( g \) is negative (directed upwards, or 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 macroscopic 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 we do not need to investigate this in detail; we only know that this is somehow reflected in the equation of state \( f(T,Y,g) = 0 \) and this dictates how \( U \) and \( C \) depend on \( g \); we need to find the equation of state and part of the information needed for this is contained in the thermal expansion coefficient \( \alpha_{\text{lin}} \); measuring this coefficient is enough 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 key to the difference between reversible and irreversible processes (changes). It allows to decide whether a given process (in an isolated system) is possible.

Let the two equilibrium states \( A \) and \( B \) of a system be connected by two paths (processes), \( \Gamma^1_{A \rightarrow B} \) and \( \Gamma^2_{A \rightarrow B} \), of which the first path is irreversible, while the second one is reversible. The Clausius inequality (54) applied to the cycle \( A \rightarrow (1) \rightarrow B \rightarrow (2) \rightarrow A \)
Figure 4: Example of a cycle executed by the perfect gas. The process $A \rightarrow C \rightarrow B$ is reversible. The isochoric process $A \rightarrow B$ can be irreversible or reversible.

can be written in the form

$$\int_{\Gamma_{A\rightarrow B}}^{1} \frac{q}{T_{ext}} + \int_{\Gamma_{B\rightarrow A}}^{2} \frac{q}{T_{ext}} < 0.$$  

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

$$\int_{\Gamma_{B\rightarrow A}}^{2} \frac{q}{T_{ext}} = -\int_{\Gamma_{A\rightarrow B}}^{2} \frac{dQ}{T} = -(S_{B} - S_{A}),$$

because 2TMDL tells us that $dQ/T = dS$. Thus in the irreversible change $A \rightarrow B$

$$\int_{\Gamma_{A\rightarrow B}}^{3} \frac{q}{T_{ext}} < \Delta S \equiv S_{B} - S_{A}. \tag{88}$$

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” $q/T_{ext}$ 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)

$$q \leq T_{ext} dS. \tag{89}$$

The Clausius inequality (88) 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 4). 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) $S_{B} - S_{A} = C_{V} \ln(T_{B}/T_{A})$, the inequality (88) applied to the transition $A \rightarrow B$ takes the form

$$\int_{A}^{B} \frac{q}{T_{ext}} \leq C_{V} \ln \frac{T_{B}}{T_{A}}.$$  

One can now contemplate different isochoric processes $A \rightarrow B$. 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 heating the gas irreversibly by connecting it thermally with a reservoir of heat of constant temperature \( T_{ext} \); the heat will flow from the reservoir to the gas provided \( T_{ext} \geq \max(T_A, T_B) = T_B \), so

\[
\int_A^B \frac{q}{T_{ext}} = \frac{1}{T_{ext}} \int_A^A \frac{Q_{A\rightarrow B}}{T_{ext}} = \frac{C_V(T_B - T_A)}{T_{ext}} \leq \frac{C_V(T_B - T_A)}{T_B}
\]

(In the last step we have taken into account that \( Q_{A\rightarrow B} \equiv U_B - U_A \) is also equal to \( W_{C\rightarrow B} = C_V(T_B - T_C) = C_V(T_B - T_A) \), as can easily be computed using the perfect gas adiabat equation \( p V^c V / C_p = \text{const.} \), and the relation \( C_p - C_V = nR \), because in the isothermal transition \( A \rightarrow C \) the energy of the perfect gas does not change). The inequality 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 the 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 \( dQ = dU \) and \( dU = C_V dT \) in the case of perfect gas)

\[
\int_A^B \frac{q}{T_{ext}} = \int_A^B \frac{dQ}{T} = \int_{T_A}^{T_B} \frac{dU(T)}{T} = C_V \int_{T_A}^{T_B} \frac{dT}{T},
\]

in which case the inequality (88) is satisfied as equality.

Returning to the general considerations, the very important case arises when the system is adiathermally isolated during the change \( A \rightarrow B \) (which means that \( q = 0 \)). The inequality (88) takes then the form

\[
\Delta S \geq 0.
\]

This is the celebrated entropy increase law (or just the entropy law):
The entropy of an isolated system can never diminish.
Clausius expressed it in the categorical German (Prussian) way: Energie der Welt ist konstant. Entropie der Welt strebt einem maximum zu.

The entropy law gives a thermodynamical criterion allowing to decide which processes of those which could occur (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. 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 \( \Delta S_2 = -q/T_2 \) 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 the heat is forbidden by the entropy law.\textsuperscript{26}

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 were the transitions from one equilibrium state to another equilibrium state as a result of strengthening (weakening of) the constraints to which the systems were subject (in the first case the constraint could be represented by an adiathermal wall which gets then replaced by a diathermal one and 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 the increase of its and of the surrounding’s temperatures. The entropy of the entire system (body and the surrounding) has 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 applicable to transitions between equilibrium states only.

For a given set of constraints to which it is subject, a thermodynamic system has only one true equilibrium state and the entropy law can be formulated as the statement (Pippard again)

**Entropy law:** It is not possible to vary the 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, we will use it (applied to the second example above) to consider briefly the role of fluctuations. When the gas fills the whole cylinder and is in equilibrium, its local density \( \rho \) seems to be uniform. It is however subject 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. 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\textsuperscript{27} (which

\textsuperscript{26}TMDL 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 “hotter” and “colder”). Logically, this requires formulating TMDL without reference to hotter and colder, that is not to base it on the Clausius formulation.

\textsuperscript{27}Of course the probability of such a fluctuation in a gas consisting of \( \sim 10^{23} \) molecules is so fantastically small that practically there is no chance to observe such a fluctuation in practice, even waiting as long as the universe’s lifetime.
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 answer is: nothing! The thermodynamical entropy which we are considering here\textsuperscript{28} stays unchanged. The continuous spectrum of fluctuations, from the minute to the largest possible ones, are part of the thermodynamic equilibrium state and not a departure from 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 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 the cylinder. Only in this way one can 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 in the smaller volume because of the wall (is subject to 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 we never talk about “coming to equilibrium” during which entropy gradually increases. Once the wall blocking the gas from expanding into the larger volume has been removed, the entropy increases and the microscopic configuration of the gas (which for a short while is still mostly in 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.

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 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 the central problem of thermodynamics is, as Callen says, 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

\textsuperscript{28}In the kinetic theory of gases one deals with a quantity called \( H \) - introduced by Boltzmann - which is sometimes identified with entropy or, more precisely, with \( -S/k_B \). The famous Boltzmann \( H \)-theorem states that \( H \) always (although this “always” also requires some qualifications) decreases with time, so \( -k_B H \) exhibits a property which makes it similar to the thermodynamical entropy. It should be stressed, however, that \( -k_B H \) 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 (whose time evolution cannot even be discussed).
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 interesting 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 taken to be only a huge fluctuation of a more or less uniform density and temperature, that is a huge fluctuation 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 compromises, that is, to rely 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 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, occurs at a negligible rate. One is then able to define entropy of physical systems of interest, apply to them the entropy law and to make predictions.

Further, 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 is not 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 to 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 therefore be tolerated.

The origin of such "paradoxes" 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 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 are 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 useless to continue it within the framework of classical thermodynamics - the proper framework for it is the statistical thermodynamics, 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 that 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. In reality there are strong indications that fluctuations cannot be used to violate the entropy law or 2TMDL. It is Maxwell himself who hypothesized a “demon” (called the Maxwell demon ever since) which could control a trapdoor connecting two vessels of gas at the same temperatures. 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 the problem 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 is 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 uniform $T$ the radiation is 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 uncertainty principle $\Delta p \Delta q \geq \hbar$ cannot be circumvented). It is therefore not true, says Pippard, that 2TMDL is only statistically true, being repeatedly violated microscopically (but never seriously, on a 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 thermodynamic systems are subjected, which determine the system’s entropy is the basis of the Callenian formulation\(^{29}\) of thermodynamics which we will discuss now. To some, giving a fourth (after the ones of Clausius, Kelvin and Caratheodory) 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 allow 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. 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 the 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 in which the work \(Z\) of change of the amount of matter in the system will (ultimately) be included: \(\Delta U = Q + W + Z\) (or \(dU = q + w + z\) if infinitesimal changes are considered). In reversible processes it will be possible to write as in the conventional approach \(w = dW\) and \(q = dQ\). If the matter transfer is done reversibly it will be possible to write \(z = dZ = \sum_{i=1}^{r} \mu_i \, dn_i\), where \(\mu_i\) is the chemical potential of the \(i\)-th component.

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

- There exists entropy \(S\) which is defined on all equilibrium states of the system and is a function of the system’s internal energy \(U\) and other globally defined parameters like its volume \(V\), magnetization \(M\), and like (denoted below collectively \(X_i\)) and of the amount of matter contained in the system and represented by the number(s) \(n_i\) of moles of its constituents. If the system is homogeneous and possesses the extensiveness property, entropy is the homogeneous function of order one of its extensive arguments.

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

\[
\left( \frac{\partial S}{\partial U} \right)_{X_j, n_i} > 0, \]

- Entropy of a system consisting of several subsystems is additive

\(^{29}\)In fact the approach we call Callenian comes directly from Gibbs and was subsequently developed by Tisza; it was written up by Callen in his textbook on thermodynamics.
• Entropy of an isolated system takes on the maximal value with respect to all possible equilibrium states which could be realized with the help of stronger constraints 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 sub-systems which are homogeneous and possesses the property of extensiveness which means that their global parameters $X_i$ (those which characterize the system as a whole, e.g. its volume, internal energy, total polarization, 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 $n_i$) scale 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, 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.\(^\text{30}\) In many cases (e.g. fluids) scaling of the dimensions of the system with 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 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 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\(^\text{31}\) forever from serious treatments of the quantum field theory!) are nevertheless systems of fields and the number of particles contained in such a system is from the point of view of thermodynamics and equilibrium statistical physics an ill defined quantity. For this reason no chemical potential can be

\(^{30}\)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).

\(^{31}\)If there is any meaning to be attached to this notion, it is rather 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.
associated with it.\footnote{In the statistical physics approach we will see that systems of fields can give rise to conserved charges - like the electric charge in quantum electrodynamics - and it is with these charges that appropriate chemical potentials must be associated.}

As far as systems possessing the extensiveness property are concerned, it is in many situations convenient to work not with extensive quantities like \( U, X \) and \( n_j \) but with molar quantities \( u, v \) etc. and the molar fractions \( x_i \) defined as

\[
u = \frac{U}{n}, \quad v = \frac{V}{n}, \quad x_i = \frac{n_i}{n}, \quad n = \sum_{i=1}^{r} n_i.
\]

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

\[
U(p, V, n_1, \ldots, n_r) = n u(p, v, x_1, \ldots, x_r),
\]

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 \( f = f(T, v, x_1, \ldots, x_r) \), its full dependence on the number \( n \) of moles can always be restored by writing

\[
F(T, V, n_1, \ldots, n_r) = n f(T, V/n, n_1/n, \ldots, n_r/n).
\]

Internal energy of the electromagnetic field must in turn take the form \( U(T, V) = V u(T) \).

Entropy of a 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:

\[
S(\lambda U, \lambda X_j, \lambda n_i) = \lambda S(U, X_j, n_i),
\]

which means that it too can be written as\footnote{The usage of various characters representing physical quantities in these lectures appears now a bit confusing. Will have to work on this, when time allows...}

\[
S(U, V, \ldots, n_i) = n S(U/n, V/n, \ldots, n_i/V) = n s(u, v, \ldots, x_i).
\]

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.

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: 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

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

One can imagine infinitely many ways\(^{34}\) 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 walls is that state among all equilibrium states which can be “fabricated” by introducing auxiliary walls in different ways, that has the largest entropy. This Callenian principle of maximum entropy pertaining to isolated systems will be later generalized to apply also to systems in different specific contacts (thermal, mechanical, material) with their surroundings.

One should note the 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 lowest $I$. 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.\(^ {35}\) This is not so in thermodynamics: the true equilibrium 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 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.\(^ {36}\) 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 or Kelvin’s formulations of 2TMDL are adopted, in the

---

\(^{34}\)One should not consider divisions of the system into subsystems so small, that surface effects could become important.

\(^{35}\)The reason for this, as well as the justification of the variational principle on which mechanics can be based is quantum mechanical.

\(^{36}\)Within thermodynamics the form of $S$ must be reconstructed from the empirical data. It is the statistical mechanics which gives the concrete, unambiguous prescription how to obtain entropy of an isolated system, if its microscopic dynamics (classical or quantum) is known.
Callenian formulation is built in it from the outset: if the 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.\textsuperscript{37} In an isolated system reversible can only be processes (changes of constraints) which happen not to increase the entropy.

The expression

\[ S = S(U, X_1, \ldots, n_1, \ldots), \]

\( S = S(U, V, n) \) in the case of a one-component fluid) is called the \textbf{fundamental relation in the entropy representation}. As will become clear, if it is known as a function of the global (extensive) variables \( U, X_i, n_j \), 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

\[ \left( \frac{\partial S}{\partial U} \right)_{V,n_i} \equiv \frac{1}{T}, \quad \left( \frac{\partial S}{\partial V} \right)_{U,n_i} \equiv \frac{p}{T}, \quad \left( \frac{\partial S}{\partial n_i} \right)_{U,V,n_i'} \equiv -\frac{\mu_i}{T}. \]

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 \( \mu_i \) (and possibly others, if the system is more complicated) \textbf{defined} in this way do indeed have the meanings of the ordinary (absolute) temperature, 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 whose energy \( U \) and the number of moles are fixed, say a fluid in a cylinder of fixed volume \( V \). Introducing an adiathermal nonmoveable 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, which are in thermal contact through a diathermal 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 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 biggest entropy. Now, entropies of the virtual equilibrium states in which the system consists of two subsystems are given (using the postulated additivity of entropy) by the formula

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

\( ^{37} \)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).
in which the departures of the virtual state energies \( U_1 \) and \( U_2 \) from the (unknown) equilibrium distribution \( U_1^{\text{eq}} \), \( U_2^{\text{eq}} \) \((U_1^{\text{eq}} + U_2^{\text{eq}} = U)\) have been parametrized with \( \delta U \) (automatically taking into account the condition \( U_1 + U_2 = U \)). Actually, if as here, the considered system is extensive, the two functions \( S_1 \) and \( S_2 \) are simply the same function \( S(\cdots) \). The extremum condition \( \delta S = 0 \) which, in view of the simplified way of seeking the maximum\(^{38}\) could simply be reduced to \( dS/d(\delta U) = 0 \), gives the equality

\[
\delta S = \left( \frac{\partial S_1}{\partial U_1} \right)_{V_1,n_1} \delta U + \left( \frac{\partial S_2}{\partial U_2} \right)_{V_2,n_2} (-\delta U) = 0.
\]

which implies that in equilibrium the parameters \( 1/T_i = (\partial S_i/\partial U_i)_{V_i,n_1} \big|_{U_i = U_i^{\text{eq}}} \) must be equal.\(^{39}\) 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,

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

that is, that \( T_1 > T_2 \). Then when the adiathermal wall is changed to a 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:

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

This means that \( \delta U \) must be positive, i.e. that the body which had higher temperature loses energy and that of lower temperature gains it.

Similarly, to show that the parameter \( p \) defined by the second of the derivatives (94) 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

\(^{38}\)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 \( q'(t) \) 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) \( \lambda \) (\( \lambda_i \)). 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 (which form only a very narrow class of possible virtual trajectories).

\(^{39}\)The assumption - which is part of the Calleanian postulates - that entropy is a monotonic function of the internal energy automatically means that that at \( U_i = U_i^{\text{eq}} \) the first derivative \( dS/d(\delta U) \) changes sign, so this point is the extremum.
\( V_i \) (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 \( n_2 \) of each, in two parts of a cylinder separated by a movable diathermal wall. The values \( U_i^{eq} \) and \( V_i^{eq} \) 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

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

that is such that

\[
\delta S = \left( \frac{\partial S_1}{\partial U_1} \right)_{V_1,n_1} \delta U + \left( \frac{\partial S_2}{\partial U_2} \right)_{V_2,n_2} \delta U + \left( \frac{\partial S_1}{\partial V_1} \right)_{U_1,n_1} \delta V + \left( \frac{\partial S_2}{\partial V_2} \right)_{U_2,n_2} (-\delta V) = 0. \tag{95}
\]

Since the departures \( \delta U \) and \( \delta V \) 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 \( 1/T_i = (\partial S_i/\partial U_i)_{V_i,n_i|U_i=U_i^{eq},V_i=V_i^{eq}} \) and of the two pressures \( p_i = T_i(\partial S_i/\partial V_i)_{U_i,n_i|U_i=U_i^{eq},V_i=V_i^{eq}} \). The reason, why the derivative of \( S \) with respect to \( V \) is denoted \( p/T \) and not \( p \), can be sought in the shock 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 5). 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 (entropy of the mechanical elements of the system is assumed to be zero)

\[
S = S_{gas}(E - \frac{1}{2} k x^2, V_0 + A x, n).
\]

Equating to zero the derivative of \( S \) with respect to \( x \) one gets the condition (the symbol \( x^{eq} \) stands for \( U = E - \frac{1}{2} k x^{eq}_2 \) and \( V = V_0 + A x^{eq} \))

\[
-k x \left( \frac{\partial S}{\partial U} \right)_{V,n|_x^{eq}} + A \left( \frac{\partial S}{\partial V} \right)_{U,n|_x^{eq}} = 0.
\]

68
Figure 5: 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.

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

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, enclosed in two adjacent parts of a cylinder (isolated as a whole) separated by a movable but adiathermal piston\footnote{A strange feature of a similar system was demonstrated in classes.} (Figure 5, 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 $V_1^{\text{eq}}$ and $V_2^{\text{eq}}$ ($V_1^{\text{eq}} + V_2^{\text{eq}} = V$) having some well defined internal energies $U_1^{\text{eq}}$ and $U_2^{\text{eq}}$ (with $U_1^{\text{eq}} + U_2^{\text{eq}} = U$). 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 nevertheless can do work on one another through the pressure they exert on one another through the movable piston. Thus as previously one can write as previous the condition (95)

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

However in this case here the variations $\delta V$ and $\delta U$ 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 $\delta U = -p_2 \delta V$ (the work on the gas number 1 equals the change $\delta V$ of its volume times the external pressure which is provided by the pressure of the gas number 2). Using this relation converts the condition (96) into

$$
\delta S = \frac{p_1^{\text{eq}} - p_2^{\text{eq}}}{T_2^{\text{eq}}} \delta V = 0 .
$$

This shows that the equilibrium position of the piston must be such as to make the pressures of both gases equal, but does not impose any condition on the equilibrium
temperatures $T_1^{eq}$ and $T_2^{eq}$. 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 arbitrary, but this neglects the role played by fluctuations. If the initial pressures of the two gases are 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 $n_1^{eq}$ and $n_2^{eq}$ of moles of the gas and of the energies $U_1^{eq}$ and $U_2^{eq}$ 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 $\delta n$ of the number of moles of the gas contained in the volume $V_1$ and $\delta U$ 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

$$
\delta S = \left( \frac{\partial S_1}{\partial U_1} \right)_{n_1,V_1} \delta U + \left( \frac{\partial S_2}{\partial U_2} \right)_{n_2,V_2} \left( -\delta U \right) + \left( \frac{\partial S_1}{\partial n_1} \right)_{n_1,V_1} \delta V + \left( \frac{\partial S_2}{\partial n_2} \right)_{n_2,V_2} \left( -\delta V \right), \quad (97)
$$

must vanish. As the departures $\delta U$ and $\delta n$ are independent, this entails the equality of the temperatures of the gases in the two subvolumes and the equality of their chemical potentials: $\mu_1 = \mu_2$. 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\textsuperscript{41} 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

\textsuperscript{41} Of course if a solid, made, say, of iron is in contact with the air, the chemical potential $\mu_{Fe}$ of the air (treated as one subsystem) is zero, similarly as is zero the chemical potential $\mu_{Air}$ of the solid. Yet no transfer of matter occurs because the iron surface should be treated as a wall impermeable to both air and 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). 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.

Since by assumption entropy of a system is a monotonic function of the internal energy $U$, the fundamental relation in the entropy representation (93) can be inverted to give

$$U = U(S, X_1, \ldots, n_1, \ldots).$$

(98)

This is called the fundamental relation in the internal energy representation. If known as a function of the system’s extensive (global) parameters it too, similarly to the relation (93), 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 $U$

$$dU = TdS - p dV + \sum_{i=1}^{r} \mu_i d n_i,$$

(99)

are, as we now know, the temperature, 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 $dW$ of the work (in the case of (99) the mechanical volume work) performed on the system and the terms $\mu_i d n_i$ 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 $dQ$ is in this approach to 2TMDL obtained almost for free!

Let us now demonstrate the fundamental role of the fundamental relation (93) 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

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

(100)

The system is clearly extensive, for $S(\lambda U, \lambda V, \lambda n) = \lambda S(U, V, n)$. Computing explicitly the derivatives (94) one finds the relations

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

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

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

71
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 $n$:

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

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

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

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

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

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

Finally, the last of the derivatives (94) 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 $p$:

$$\mu(T, v) = -a^{3/2} T^{3/2} \left( \frac{V}{n} \right)^{1/2} = -\frac{a^{3} T^{3} p}{27 a^{3} n V} = \mu(T, p).$$  

It is left for the students to check that the same results can be derived from the fundamental relation (98) in the energy representation which in this case reads

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

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.

**Digression. Black hole entropy (example of a nonextensive system).**

A 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 $L$. It is known that $M, Q$ and $L$ are the only possible quantities which can characterize a black hole (famous saying that “black holes have no hair”).

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 $L$, cannot tell how much entropy the black hole has swallowed. This has led J.D. Bekenstein, 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 $L$. This was prompted by a theorem of S. Hawking which says that according to the laws of classical physics, 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 rotational energy can 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

---

42 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.

43 Just a year ago, or so, I saw his obituary in the CERN Courier.

73
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:

\[
S = \frac{1}{4} k_B \frac{A}{\ell_{Pl}^2}.
\]

(101)

\(k_B\) is the Boltzmann constant and \(\ell_{Pl} \equiv \sqrt{\hbar G/c^3}\) is the Planck length
\((G = 6.67 \times 10^{-11}\) m\(^3\)kg\(^{-1}\)sec\(^{-2}\) is the gravitational constant; \(\ell_{Pl} = 1.6 \times 10^{-35}\) 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 \(A\) of
the Schwarzschild black hole \((Q = 0, L = 0)\) is very simple:

\[
A_{Sch} = 4\pi R_g^2 = 16\pi \frac{G^2}{c^4} M^2 = 16\pi \frac{\hbar G}{c^3} \frac{M^2}{M_{Pl}^2} = 16\pi \ell^2_{Pl} \frac{M^2}{M_{Pl}^2},
\]

(102)

here \(M_{Pl} = \sqrt{\hbar c/G} = 2.2 \times 10^{-8}\)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
\(1/2mv_{II}^2 - GMm/r = 0\) 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. Thermodynamics
of the Schwarzschild black hole is very simple:

\[
S = 4\pi k_B \frac{M^2}{M_{Pl}^2},
\]

\[
U = Mc^2.
\]

(103)

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 equation

\[
S(M) = 4\pi k_B \frac{U^2}{M_{Pl}^2 c^4},
\]

(104)

plays the role of the fundamental relation (93) of the Schwarschild black hole thermody-
namics. It is to be noted that in this case \(S\) is not a homogeneous first order function
of its only global parameter \(U\). The simple scaling is violated by the presence of the
fundamental mass (or length scale) \(M_{Pl}\) (or \(\ell_{Pl}\)). Applying to \(104\) the standard formula
\(dS = dU/T\) one gets

\[
\frac{1}{T} = \frac{dS}{dU} = \frac{8\pi k_B}{M_{Pl}^2 c^4} U,
\]

(105)

or

\[
U = \frac{M_{Pl}^2 c^4}{8\pi k_B T},
\]

(106)
The larger 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

\[
C = T \frac{dS(U(T))}{dT} = \frac{dU}{dT} = -\frac{k_B}{8\pi} \left( \frac{M_{Pl} c^2}{k_B T} \right)^2.
\]  

(107)

If a black hole has nonzero angular momentum \( L \) 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 a coupling to \( L \) or to \( Q \) appropriate external agents. Such reversible works can be continued until \( L = 0, Q = 0 \). 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 \( Q^2 / hc \) is dimensionless - \( e^2 / hc = \sigma_{EM} \) as every high energy physicist knows)

\[
\tilde{Q}^2 \equiv \frac{Q^2}{G} = \frac{Q^2 h c}{G} = \frac{Q^2}{h c} M_{Pl}^2,
\]

\[
a^2 = \frac{c^2}{G^2 M^2} \frac{L^2}{M^2} = \frac{L^2}{M^2} \left( \frac{h c}{G} \right)^2 \frac{1}{M^2} = \frac{L^2}{M^2} M_{Pl}^4,
\]

and reads

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

(108)

This formula together with (101) and the identification \( U = M c^2 \), gives the fundamental Callenian relation of the most general black hole thermodynamics:

\[
S = S(U, L, Q).
\]  

(109)

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

\[
dS = \frac{1}{T} dU - \frac{\phi}{T} dQ - \frac{1}{T} \Omega \cdot dL,
\]

(110)

which leads to the form

\[
dU = TdT + \phi dQ + \Omega \cdot dL,
\]

(111)

of \( 1TMDL \) as applied to reversible changes of the black hole. \( \phi \) is here the electrostatic potential and \( \Omega \) 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 (108), 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 $C_\phi$ of the Reisner-Nordström ($\mathbf{L} = 0$) black hole or $C_\Omega$ of the Kerr ($Q = 0$) one.

The formula for the Schwarzschild black hole temperature can be written in terms of the so-called surface gravity $\kappa$

$$
T = \frac{\hbar}{2\pi k_B c} \frac{M^2 c^3}{4\hbar M} = \frac{\hbar}{2\pi k_B c} \left( \frac{c^4}{4GM} \right) \equiv \frac{\hbar}{2\pi k_B c} \kappa, \quad (112)
$$

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 $R_g$.

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.\(^{44}\) 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 (101) 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\(^{46}\) discovered (in 1974) that it radiates just as does a body at temperature $T$ given by (112). 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.

\(^{44}\)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.

\(^{46}\)It is these considerations which allowed Hawking to fix the coefficient in the Bekenstein formula (101).
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. These considerations will lead us to the formulation of the principle which determines the equilibrium states of systems open to their surroundings, generalizing thereby the Callenian maximum entropy principle which applies to isolated systems only. Also the general conditions of stability of thermodynamic systems will be considered.

It is easy to show 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

$$S_A = S_A(\lambda_A U_A, \xi_A V_A, \ldots), \quad S_B = S_B(\lambda_B U_B, \xi_B V_B, \ldots),$$

that is, we consider different extensive parameters of the two system scaled up by the respective arbitrary factors $\lambda$, $\xi$, etc. We now imagine that the two systems have been brought into a thermal contact and a new equilibrium state has been reached. By the entropy law, the entropy of this new equilibrium state, $S_{A+B}(\lambda_A U_A + \lambda_B U_B, \xi_A V_A + \xi_B V_B, \ldots)$ cannot be smaller than the sum of the entropies $S_A$ and $S_B$:

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

irrespectively of the values (positive) of the factors $\lambda$, $\xi$, 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 $S(\cdot, \cdot, \ldots)$ but evaluated at different values of its arguments. If in addition one sets $\lambda_A = \lambda$, $\lambda_B = 1 - \lambda$, $\xi_A = \lambda$, $\xi_B = 1 - \lambda$, etc. and considers $V_A = V_B = V$, etc., one obtains the inequality

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

which just means concavity upwards of the entropy $S(\cdot, \cdot, \ldots)$ of the system as a function of its first (energy) argument (Figure 6). In the same way one can demonstrate its concavity as a function of its volume argument and all other extensive parameters.

Concavity upwards of the entropy as a function of the internal energy is equivalent to the statement that

$$\left( \frac{\partial^2 S}{\partial U^2} \right)_{V, \ldots} < 0.$$  \hfill (114)

Because of the first of the equalities (94) this amounts to

$$\left( \frac{\partial}{\partial U} \frac{1}{T} \right)_{V, \ldots} = -\frac{1}{T^2} \left( \frac{\partial U}{\partial T} \right)^{-1}_{V, \ldots} \equiv -\frac{1}{T^2 C_V} < 0,$$  \hfill (115)

77
Figure 6: Entropy of extensive systems is a concave upwards function of internal energy. The dashed line shows the function $\lambda S(U_A) + (1 - \lambda)S(U_B)$. Everywhere between $U_A$ and $U_B$ the value of $S(\lambda U_A + (1 - \lambda)U_B)$ is above this line.

which, in view of the positivity of the temperature (which is a direct consequence of the third Callenian postulate), in turn 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 entropy on the set of all virtual equilibrium states (that can be fabricated with the help of constraints stronger than the ones the system is actually subjected to) at the fixed value of the internal energy $U$ (and fixed values of other global parameters) - the **maximum entropy principle** - is equivalent to the statement that the equilibrium state minimizes internal energy on the set of all virtual states at the fixed value of the system’s entropy $S$ - this will be called the **minimum energy principle**. This second statement sounds a little bit more abstract - it is easier to imagine keeping internal energy fixed (and distributing it among subsystems) than to imagine fixed entropy and varying 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 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 of energy $U$ 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 $S_0$, logically this would be possible!) extracting from it the work $\dot{W} = U_0 - U$. The obtained in this way energy could be then put back irreversibly into the system 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 constant energy because there would be a state (possibly with stronger constraints) of the same energy and higher entropy.

The proof in the reverse direction (that the minimum energy principle implies the maximum entropy principle) goes similarly: we assume that the 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 the system is subjected to stronger constraints); the system in this higher entropy state could be then brought into a contact 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 be, therefore, the state of lowest possible energy (for fixed \(S_0\)).

The mathematical reasoning requires specifying a parameter, say \(x\) (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

\[
\left( \frac{\partial S}{\partial x} \right)_{U,...,x=0} = 0, \quad \left( \frac{\partial^2 S}{\partial x^2} \right)_{U,...,x=0} < 0. \tag{116}
\]

Using the “shocking” relation one can write

\[
\left( \frac{\partial U}{\partial x} \right)_{S,...,x=0} = - \left( \frac{\partial U}{\partial S} \right)_{x,...} \left( \frac{\partial S}{\partial x} \right)_{U,...,x=0} = 0. \tag{117}
\]

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

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

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

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

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

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

which shows that if \(S(U, x, \ldots)\) has a maximum at \(x = 0\), then \(U(S, x, \ldots)\) has a minimum. The proof can be easily 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.
The conclusion is that with respect to isolated systems entropy and internal energy play the role 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 mechanical work which can be extracted from a given thermodynamical system by changing appropriately the constraints to which the system is subjected, thereby allowing it to reach another equilibrium state. The reverse 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 is as a whole adiathermally isolated. With the environment (the rest of the world) it can only exchange mechanical work. 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 deliver some work; since the system is not completely isolated\(^{40}\) (it can exchange work with the environment) the final equilibrium state and the work extracted from the system depend on the process by which the new equilibrium 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.\(^{47}\) Of course, it can vary during the process. In addition, some subsystems of the considered system may be subsidiary in the sense that their initial and final states are the same.

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

\[
\bar{W} = U_{in} - U_{fin}.
\]

The energy \(U_{in}\) is fixed but the final state and, therefore, \(U_{fin}\) 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 biggest 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

\(^{40}\)Recall that if the system were completely isolated, its final state would be uniquely determined by the new constrains.

\(^{47}\)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 expand 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 environment is negligible (or zero) so that expanding the system does not do any work.
have, hence, the same final temperature. If the two bodies are brought into the 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 the preceding Lecture) and no work at all will be extracted from the system. Assuming for simplicity that the heat capacities $C$ at constant volume of the two bodies are equal and (to make things simple) independent of the temperature, the final common temperature $T_{\text{fin}}$ will be determined by 1TMDL:

$$U_{\text{in}} = (U_0 + CT_1) + (U_0 + CT_2) = (U_0 + CT_{\text{fin}}) + (U_0 + CT_{\text{fin}}) = U_{\text{fin}}, \quad (119)$$

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

$$\Delta S = 2C \ln \frac{T_{\text{fin}}}{T_0} - C \ln \frac{T_1}{T_0} - C \ln \frac{T_2}{T_0} = C \ln \frac{T_{\text{fin}}^2}{T_1T_2} = C \ln \frac{(T_1 + T_2)^2}{4T_1T_2} > 0. \quad (120)$$

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 formula $\Delta S = C \ln(T_{\text{fin}}^2/T_1T_2) = 0$ which will determine the final temperature $T_{\text{fin}} = \sqrt{T_1T_2}$. By 1TMDL the work extracted from the system in this case will be

$$\dot{W} = U_{\text{in}} - U_{\text{fin}} = 2U_0 + C(T_1 + T_2) - 2 \left( U_0 + C\sqrt{T_1T_2} \right)
= C \left( T_1 + T_2 - 2\sqrt{T_1T_2} \right) = C \left( \sqrt{T_2} - \sqrt{T_1} \right)^2. \quad (121)$$

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 upper (higher temperature) heat reservoir and the body at $T_1$ as the lower one. The heat capacity of the working substance, which is a subsidiary element of the system, should be small compared to $C$ to ensure that the bodies stay practically in equilibrium when the heat is transferred between them and the engine. 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 upper reservoir and raises the one of the lower reservoir, so that finally they will equalize. If the actual (during the working of the engine) temperature of the higher reservoir is $T_+ \ (T_{\text{fin}} \leq T_+ \leq T_2)$ and that of the lower one is $T_- \ (T_1 \leq T_- \leq T_{\text{fin}})$ then (upon using the actual efficiency $\eta(T_-, T_+)$ of the Carnot cycle working between the temperatures $T_-$ and $T_+$)

$$d\dot{W} = \eta(T_-, T_+)dQ_2 = \left( 1 - \frac{T_+}{T_+} \right) \left( -C dT_+ \right),$$

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

$$\frac{dQ_2}{T_+} + \frac{dQ_1}{T_-} = -C \left( \frac{dT_+}{T_+} + \frac{dT_-}{T_-} \right) = 0,$$

$^{48}$The most general case is treated in one of the homework Problems.

$^{49}$As said, one could also consider the heat capacities at constant internal pressure if the external pressure is zero.
of constancy of the entropy ($dQ_2 = -CdT_+ > 0$ and $dQ_1 = -CdT_- < 0$ are the heats taken by the engine from the two bodies) which upon integration with the obvious initial conditions yields $T_- = T_1T_2/T_+$. Integrating then $dW$ from $T_2$ to $T_{\text{fin}}$ gives the same result (121).

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

A more general situations is when a system, which can be composed of several sub-systems, can exchange with the environment 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 environment as consisting of a reversible 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 was extracted from or added to it, the reversible volume source (RVS) - a very large system in equilibrium (whose entropy stays always constant and need not be taken into account) at invariable pressure $p_0$ and, finally, a source of matter\(^{50}\) (RMS) in equilibrium (whose entropy also stays always constant and need not be taken into account - recall that the chemical potential is defined by (94) 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 $\mu_0$ (this general setting is schematically illustrated in Figure 7). Owing to the constraints (internal and external, separating the system from its

---

\(^{50}\)If there is more than one matter component which can be exchanged with the environment, we introduce one such source per component.
surrounding), the system consisting of several subsystems is initially in equilibrium (the individual subsystems can have 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.\(^{51}\) In the process of reaching the new equilibrium state the system may exchange heat with the RHS. We also assume that any change in 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.

By 1TMDL, the work extracted from the system and its environment represented by RHS, RVS and RMS is

\[
\dot{W} = -(\Delta U + \Delta U_{\text{RHS}} + \Delta U_{\text{RVS}} + \Delta U_{\text{RMS}}),
\]

where \(-\Delta U = U^{\text{in}} - U^{\text{fin}}\) etc. But since the reservoirs stay in equilibrium, \(\Delta U_{\text{RHS}} = T_0 \Delta S_{\text{RHS}}, \Delta U_{\text{RVS}} = -p_0 \Delta V_{\text{RHS}}\) and \(\Delta U_{\text{RMS}} = \mu_0 \Delta n_{\text{RMS}}\). Furthermore, by assumption \(\Delta V_{\text{RHS}} = -\Delta V\) and \(\Delta n_{\text{RMS}} = -\Delta n\) (the total volume and the total amount of matter in the system and its environment remain constant\(^{52}\))

\[
\dot{W} = -(\Delta U + T_0 \Delta S_{\text{RHS}} + p_0 \Delta V - \mu_0 \Delta n).
\]

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 that \(\Delta S_{\text{RHS}} + \Delta S \geq 0\), or that

\[-T_0 \Delta S_{\text{RHS}} \leq T_0 \Delta S.\]

Therefore the work \(\dot{W}\) which can be extracted from the system accomplishing a change (specified by \(\Delta U, \Delta V\) and \(\Delta n\)) in the given environment (characterized by \(T_0, p_0\) and \(\mu_0\)) is limited form above:

\[
\dot{W} \leq -\Delta (U - T_0 S + p_0 V - \mu_0 n) = -\Delta A.
\]

The quantity \(A \equiv U - T_0 S + p_0 V - \mu_0 n\) is called **availability of the system**. Contrary to this name, it is the property of the system and its environment. Recall also that in

\(^{51}\)In the setting considered here also the same final state can be reached through different processes.

\(^{52}\)If there were more matter constituents in the system (and correspondingly more matter sources representing the environment) and chemical reactions were allowed to occur in the system (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 and occurring due to the exchanges with the matter sources in the environment (the changes of numbers of moles due to reactions would then be part of specification of the process by which final equilibrium state is reached); no problem in accounting for chemical reactions arises, however, if the system does not exchange matter with the environment and the term \(\Delta U_{\text{RMS}}\) in (122) is absent (formally one can set \(\mu_0\) in all further formulae).
general the quantities $U$, $V$ and $S$ pertaining in (123) to the system are the sums of energies, volumes and entropies of its different parts (subsystems).

If the change of the system is infinitesimal, replacing $\Delta U$ by $dU = TdS - p\,dV + \mu\,dn$ one gets$^{53}$

$$\bar{w} \leq -(T - T_0)dS + (p - p_0)dV - (\mu - \mu_0)dn.$$  \hfill (124)

If the change is reversible, $-TdS$ is the heat $dQ$ lost by the system; its conversion into work in the environment at temperature $T_0$ (e.g. by running a suitable Carnot engine) yields the useful work equal $\eta(T_0, T)dQ = -(T - T_0)dS$. 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 RWS) and the work done by the system on RWS is just $(p - p_0)dV$. Finally, to make the change of matter content of the system by $-\Delta n$ reversible requires additional chemical potential $\mu - \mu_0$ to be provided by the RWS$^{54}$ and the work done by the system on RWS due to the matter flow is $-(\mu - \mu_0)dn$.

Analogously, the minimal work $W$ needed to accomplish a given change (strengthening its constraints and characterized by $\Delta U$, $\Delta V$ and $\Delta n$) of the system in the given

\[ \bar{w} \leq -\sum_a [(T_a - T_0)dS_a + (p_a - p_0)dV_a - (\mu_a - \mu_0)dn_a]. \]

$^{53}$We restrict here ourselves to a system consisting of a single homogeneous body or to a situation in which all parts of the system have the same temperature $T$ and pressure $p$; in the general case of a system consisting of several homogeneous bodies each of which could have a different temperature, pressure and chemical potential the formula (124) would take the form

$^{54}$Here we enlarge a bit the meaning of the work source, allowing it to change its energy also in the form of matter flow. If one keeps the original definition of RWS, a reversible change of matter content of the system requires its chemical potential $\mu$ to be equal to $\mu_0$ and the last term in (124) 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 (124) is valid only as the inequality.

84
surrounding is bounded from below
\[ W \geq \Delta (U - T_0 S + p_0 V - \mu_0 n) = \Delta A. \] (125)

The minimal work \( W_{\text{min}} \) needed to accomplish the change is given an interpretation in Figure 8 which sketches the dependence of the total entropy of the system and of its
evironment on the total energy with weaker (solid line) and stronger constraints (dashed line); the horizontal dotted line shows the change of the total energy needed to strengthen
the constraints at fixed entropy (it is thus equal \( W_{\text{min}} \)); the vertical dotted line shows
the increase \( \Delta S_{\text{tot}} > 0 \) of the total entropy if the stronger constraints are replaced by
the weaker ones and the system and its environment attain equilibrium without any
intervention from without (i.e. from RWS). If the system is very small compared to
the environment (treated here as real and not as an infinitely large) then \( W_{\text{min}} \ll U_{\text{tot}} \)
and the increase of the total entropy (in the spontaneous attaining the equilibrium after
weakening the constraints) can be written as (here \( U_{\text{tot}} \) and \( S_{\text{tot}} \) stand for the total energy
and entropy of the system and of the surrounding)
\[ \Delta S_{\text{tot}} \approx \left( \frac{\partial S_{\text{tot}}}{\partial U_{\text{tot}}} \right)_{V_{\text{tot}}, n_{\text{tot}}} \frac{W_{\text{min}}}{T_0} \left( \Delta U - T_0 \Delta S + p_0 \Delta V - \mu_0 \Delta n \right). \] (126)

This formula quantifies by how much the total entropy of the system and of its environ-
ment (treated together as an isolated supersystem) with given constraints differs from
the maximal entropy obtained when the system is in equilibrium with its environment
(without walls blocking the direct contact between them).

The formula (123) is very general and applies to most of the situations in which the
system exchanges heat with the environment. A few special cases deserve to be discussed:

- The formula (123) has been derived assuming that the system does not exchange
heat with the environment (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 heat
exchange cannot occur between its different subsystems) the inequality (123) be-
comes equality: the work done on the RWS in the process in which the system
reaches a given final state equals simply the total change of energy of the system, of
the volume source (RVS) and of the matter source (RMS). If however, as in the pre-
viously discussed situation, one is interested in the system change specified only by
the final constraints (and not by the final state), the work \( \dot{W} = -\Delta (U + p_0 V - \mu_0 n) \)
done on RWS is maximal when the entropy of the system does not increase; this
generalizes the previous considerations by allowing for a change of the system’s to-
tal volume and matter content due to its exchange with the environment (RMS).
If the matter content of the system changes as a result of chemical reactions (and
there is no exchange of matter with the environment) the work done on RWS is just
\( \dot{W} = -\Delta (U + p_0 V) \) (again it is maximal for that final state which correspond 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

85
of the system are the same and equal to the pressure \( p_0 \) of the environment (e.g. the chemical reactions in the system occur under the pressure of the environment), then the work \( \bar{W} \) is given by the change of enthalpy of the system

\[
\bar{W} = -\Delta(U + pV) = -\Delta H.
\]  

\( \text{• If the system exchanges with the environment only heat (chemical reactions occurring within the system are therefore allowed) and the final total volume of all its parts is identical with the initial one,}^{55} \text{ then } \bar{W} \leq -\Delta(U - T_0S). \text{ 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 \text{ with } T, \\
\bar{W}_{\text{max}} = -\Delta(U - TS) = -\Delta F. \]

In this case the maximal work is given by the change of the Helmholtz free energy of the system.

\( \text{• If the system exchanges heat and its volume changes but there is no matter exchange with RMS (chemical reactions within the system are therefore allowed), } \bar{W} \leq -\Delta(U - T_0S + p_0V). \text{ If in addition the initial and final states of all parts of the system are in thermal and mechanical equilibrium with the environment } (T = T_0 \text{ and } p = p_0) \text{ then} \\
\bar{W}_{\text{max}} = -\Delta(U - TS + pV) = -\Delta G. \]

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 environment not changing the 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

\[
\bar{W}_{\text{max}} = -\Delta(U - T_0S). 
\]

Using the formulae (which by now should already be well-known!)

\[
U(T, V) = C_V T + \text{const}, \quad S(T, V) = C_V \ln T + nR \ln (V/n) + \text{const},
\]

(for simplicity constant heat capacity of the gas has been assumed) we readily get the answer

\[
\bar{W}_{\text{max}} = C_V (T - T_0) + T_0 C_V \ln \frac{T_0}{T}.
\]

\(^{55}\)This can be relaxed if the pressure \( p_0 \) of the environment vanishes.
The process allowing to extract this work can of course be realized with the help of an infinitesimal Carnot engine. One can also ask the question what maximal work can be obtained by cooling the gas to the temperature \( T_0 \) of the environment keeping it at constant pressure \( p_0 \) (or, more generally, if the initial and final gas pressures are equal \( p_0 \)), if it initially had the temperature \( T \). In this case one has to consider also the volume changes and compute the entropy change with temperature at constant pressure

\[
\Delta V = \frac{nR}{p_0} (T_0 - T), \quad \Delta S = C_p \ln \frac{T_0}{T},
\]

and

\[
W_{\text{max}} = C_V (T - T_0) + T_0 C_p \ln \frac{T_0}{T} - nR(T_0 - T) = C_p (T - T_0) + T_0 C_p \ln \frac{T_0}{T}.
\]

**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) generalizing thereby the Callenian maximum entropy principle (which applies to isolated systems only) and to discuss stability conditions which thermodynamical systems should satisfy in typical situations. Some exceptions will be also mentioned.

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 \( \mu_0 \). Without the external constraints the system might not be in equilibrium with the surrounding. If it is possible to change the constraints, to which the system is subjected, so that its availability decreases, a useful (i.e. positive) work can in principle be extracted from the system through an appropriate reversible process (the formula (123) 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 is stable: any change of constraints requires supplying to it a positive work from outside (the minimal such work needed is of course equal \( W_{\text{min}} = + \Delta A > 0 \)). It follows that in a given surrounding represented by \( T_0 \), \( p_0 \) and \( \mu_0 \), 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.

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 true - see the example of the equilibrium of a liquid drop with its vapour discussed 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, so that the
variables \( n \) or \( n_1, \ldots, n_r \) remain fixed) the change of the availability corresponding to arbitrary departures \( \Delta S \) and \( \Delta V \) of the system's entropy and volume from the (supposed) equilibrium state (the derivatives are taken at equilibrium values of \( S \) and \( V \)):

\[
\Delta A = A^{eq} + \left[ \left( \frac{\partial U}{\partial S} \right)_{V,\ldots} - T_0 \right] \Delta S + \left[ \left( \frac{\partial U}{\partial V} \right)_{S,\ldots} + p_0 \right] \Delta V + \frac{1}{2} \frac{\partial^2 U}{\partial S^2} (\Delta S)^2 + \frac{1}{2} \frac{\partial^2 U}{\partial V^2} (\Delta V)^2 + \frac{\partial^2 U}{\partial V \partial S} \Delta V \Delta S + \ldots
\]

(130)

It follows that the minimum of \( A \) 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, if \( T = T_0 \) 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:

\[
\left( \frac{\partial^2 U}{\partial S^2} \right)_{V,\ldots} = \left( \frac{\partial T}{\partial S} \right)_{V,\ldots} = \frac{T}{C_V} > 0, \quad \left( \frac{\partial^2 U}{\partial V^2} \right)_{S,\ldots} = - \left( \frac{\partial p}{\partial V} \right)_{S,\ldots} = \frac{1}{V k_s} > 0, \quad \frac{C_V}{k_T} > 0.
\]

and (writing \( \left( \frac{\partial^2 U}{\partial S \partial V} \right) \) in one corner of the matrix of the second derivatives as \( - \left( \frac{\partial p}{\partial S} \right) \) and as \( \left( \frac{\partial T}{\partial V} \right) \) in the other)

\[
- \left( \frac{\partial p}{\partial V} \right)_S \left( \frac{\partial T}{\partial S} \right)_V + \left( \frac{\partial p}{\partial S} \right)_V \left( \frac{\partial T}{\partial V} \right)_S = \frac{\partial (p,T)}{\partial (S,V)} = \frac{\partial (p,T)}{\partial (V,T)} \frac{\partial (V,T)}{\partial (S,V)} = \frac{T}{V C_V k_T} > 0.
\]

Thus the stability of the system requires strict positivity of its heat capacity \( C_V \) at constant volume and of its adiathermal, \( k_s \), as well as isothermal, \( k_T \), compressibilities.\(^{56}\)

Positivity of \( k_T \) is obviously required by mechanical stability of the system (were it negative, the system would spontaneously compress itself at constant temperature). Thus under normal circumstances the availability takes its minimal value at \( T = T_0 \) and

\(^{56}\) 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 \( p \); furthermore, 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).
$p = p_0$, as expected, and stability is ensured by positive (and not infinite) values of $C_V$, $k_S$ and $k_T$. Similar stability conditions can be derived for other simple thermodynamic systems like wires or rubber bands subjected to stretchings (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 $(\partial K/\partial L)_T > 0$) or magnetic materials in external magnetic fields ($C_M > 0$, $(\partial H_0/\partial M)_S > 0$ and $(\partial H_0/\partial M)_T > 0$).

There are however 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.\(^\text{57}\) There is then a direction in the $(S, V)$ space in which the second order term in the expansion of $A$ does not grow. This occurs 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). 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\(^\text{58}\) of departures $\Delta S$ and $\Delta V$ 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 $C_V$ and $k_S$ can never become infinite if $A$ can be expanded in powers of the departures $\Delta S$ and $\Delta V$ 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 $\alpha$ and $\beta$ in equilibrium with the surrounding at $T_0$ and $p_0$ which are such that both phases may coexist with an arbitrary proportion $n_\alpha / n_\beta$ ($n_\alpha + n_\beta = n$) 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 $n_\alpha / n_\beta$ is thereby affected. 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. The valley must be in the null direction direction of the quadratic form of second derivatives in (130):

$$
\begin{bmatrix}
(\partial T/\partial S)_V & (\partial T/\partial V)_S \\
(\partial T/\partial V)_S & -(\partial p/\partial V)_V
\end{bmatrix}
\begin{bmatrix}
\Delta S \\
\Delta V
\end{bmatrix} =
\begin{bmatrix}
0 \\
0
\end{bmatrix}.
$$

(131)

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

---

\(^{57}\)Recall the home Problems in which one shows that $C_p = C_V + TV\alpha_p^2/k_T$ and $k_S = (C_V/C_p)k_T$. From these relations it follows that $C_p \geq C_V$ (the equality requires either $\alpha_p = 0$ or infinite $k_T$) and, consequently, $k_T \geq k_S$.

\(^{58}\)Tetra-form like the tetradrachm - a Greek coin, an artefact from antiquity - or the tetrarchy - the system of governing the Roman empire established by Diokletianus.
value of \( k_T \) one gets that the valley direction is such that

\[
\frac{\Delta S}{\Delta V} = - \left( \frac{\partial p}{\partial V} \right)_S = \left( \frac{\partial S}{\partial V} \right)_T = \left( \frac{\partial P}{\partial T} \right)_V,
\]

(132)

(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

\[
\Delta S = s_\alpha \Delta n_\alpha + s_\beta \Delta n_\beta, \quad \Delta V = v_\alpha \Delta n_\alpha + v_\beta \Delta n_\beta,
\]

but since \( n_\alpha + n_\beta = n \) is fixed, \( \Delta n_\alpha = -\Delta n_\beta \) and therefore (132) takes the form\(^59\)

\[
\frac{dp}{dT} = \frac{s_\alpha - s_\beta}{v_\alpha - v_\beta},
\]

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 nor the denominator on the right hand side 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 finite (though \( k_S \) and \( C_p \) are infinite). On the melting curve of \( 2 \)He\(^3\) (the lighter isotope of Helium) there is 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 (131) 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 (\( \text{H}_2\text{O} \) has many different phases), which in the analogous way would maen vanishing of \((\partial^2 U/\partial S^2)_V = 0\), that is \( C_V = \infty \). This shows that the finiteness of \( C_V \) and \( k_S \) are not absolute thermodynamic requirements (but exceptions are rare).

Special cases of the general stability criterion (determining in the variational way the equilibrium states of thermodynamic systems in different conditions) formulated above deserve consideration.

- If the system is isolated, \( U + p_0V \) stays constant\(^60\) and the condition of minimal value of the availability \( A = U + p_0V - T_0S \) is equivalent to the condition of maximum entropy discussed before.

- If the volume of the system is fixed (there is no contact with RVM) and \( p - p_0 \) in the expansion (130) is indeterminate. 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 \( F(T, V) \) at constant \( V \) and constant \( T = T_0 \).

\(^59\)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.

\(^60\)If it 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 \( U + p_0V \) is conserved, because \( U + U_{\text{RVS}} \) is.
• If the system is in thermal and mechanical contact with its equilibrium at \( T_0 \) and \( p_0 \), and if in equilibrium all its parts have 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 it states are determined by minimizing the Gibbs function \( G = U - TS + pV \) at fixed \( T = T_0 \) and \( p = p_0 \).

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 a thermal contact with the environment (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 suitable wall is

\[
A = n_l(u_l - T_0 s_l + p_0 v_l) + n_v(u_v - T_0 s_v + p_0 v_v),
\]

(133)

where \( n_l \) (\( n_v \)) is the number of moles of the liquid (vapour) in the vessel. Obviously, \( n_l + n_v = n \) and all quantities \( u_{l,v} \) 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 and from the previous considerations we know that these must be equal to \( T_0 \) and \( p_0 \), respectively. The equilibrium values \( n_i^{eq} \) and \( n_v^{eq} \) (also satisfying the relation \( n_l^{eq} + n_v^{eq} = n \)) minimize \( A \) (or equivalently, the Gibbs function, \( G(T_0, p_0, n_l, n_v) \) at constant \( T_0 \) and \( p_0 \)). That is \( \delta A = 0 \) at \( n_l = n_l^{eq} \). From this we find the equilibrium condition

\[
u_l(T_0, p_0) - T_0 s_l(T_0, p_0) + p_0 v_l(T_0, p_0) = u_v(T_0, p_0) - T_0 s_v(T_0, p_0) + p_0 v_v(T_0, p_0),\]

that is, \( g_l(T_0, p_0) = g_v(T_0, p_0) \) or, as we will see shortly, \( \mu_l(T_0, p_0) = \mu_v(T_0, p_0) \). One should also notice that in this case the second derivative of \( A \) with respect to \( n_l \) vanishes identically - the equilibrium is of the neutral nature (the value of \( n_l \) is not fixed by the equilibrium condition); this has already been discussed in this Lecture.

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 \( U = n_l u_l + n_v u_v \) and fixed volume \( V = n_l v_l + n_v v_v \). Minimization of the system’s availability (133) reduces now to maximizing the total entropy \( S = n_l s_l + n_v s_v \), because \( U \) and \( V \) 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 bath). However to seek the extremum of \( S \) respecting the constraints we use the method of the Lagrange multipliers\(^{61}\) and equate to zero the

\(^{61}\)Of course, were we given the explicit forms of \( s_l(u_l, v_l) \) and \( s_v(u_v, v_v) \), we could write down the total entropy \( S \) as an explicit function \( S(U, V, n, u_l, n_l) \), say, and maximize it directly treating \( 0 \leq u_l \leq U/n \) and \( 0 \leq v_l \leq V/n \) as independent variables; in equilibrium both subsystems should have the same \( T \) and \( p \) and, therefore, all the molar quantities \( u_l, v_l, s_l \) and \( u_v, v_v, s_v \) are parametrized by the (unknown yet) temperature \( T \) and pressure \( p \); these two variables can be conveniently traded for \( u_l \) and \( v_l \).
variation of the auxiliary function in which the parameters $T$ and $p$ are the Lagrange multipliers\(^\text{62}\)

$$
\delta(U - TS + pV) = 0.
$$

The virtual equilibrium states which should be considered now 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

$$
n_1(\delta u_1 - T\delta s_1 + p\delta v_1) + n_\nu(\delta u_\nu - T\delta s_\nu + p\delta v_\nu) + \delta n_1(u_1 - Ts_1 + pv_1) + \delta n_\nu(u_\nu - Ts_\nu + p\nu_\nu) = 0.
$$

We know however, that the two Lagrange multipliers $T$ and $p$ can be chosen in such a way as to kill the coefficients of $n_1$ and $n_\nu$. This is possible, because $\delta u_1$, $\delta v_1$ and $\delta s_1$ are changes of the parameters between two equilibrium states of one mole of 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 liquid; the same pertains also to the changes $\delta u_\nu$, $\delta v_\nu$ and $\delta s_\nu$. Thus in this way, the Lagrange multipliers acquire the proper meaning of the system’s temperature and pressure and the equilibrium condition is

$$
\delta n_1(u_1 - Ts_1 + pv_1)|_{T, p} + \delta n_\nu(u_\nu - Ts_\nu + p\nu_\nu)|_{T, p} = 0,
$$

that is, because $\delta n_1 = -\delta n_\nu$,

$$
g_1(T, p) = g_\nu(T, p),
$$

which of course must be solved together with the conditions

$$
n_1u_1(T, p) + n_\nu u_\nu(T, p) = U, \quad n_1v_1(T, p) + n_\nu v_\nu(T, p) = V,
$$

to yield the equilibrium parameters $T$, $p$, $n_1$ and $n_\nu = n - n_1$. The condition (134) is, as we will discuss, the same as $\mu_1(T, p) = \mu_\nu(T, p)$ (because there is only a single component).

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 environment. In this case minimizing $A$ reduces, because $V$ is fixed, to minimizing $U - T_0S$, or because the external temperature $T_0$ is also the temperature of the system, to minimizing\(^\text{63}\)

$$
F(T_0, V, n_1, n_\nu) = n_1f_1(T_0, v_1) + n_\nu f_\nu(T_0, v_\nu)
$$

\(^{62}\)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 $\lambda_1$ and $\lambda_2$, but this is clearly equivalent to $\delta(U - TS + pV) = 0$.

\(^{63}\)Again, since in equilibrium both phases must have the same pressure, $v_1$ and $v_\nu$ are not independent and if the explicit formulæ for $v_1 = v_1(T_0, p)$ nd $v_\nu = v_\nu(T_0, p)$ were given, as well as the explicit forms of $f_1(T_0, v_1)$ and $f_\nu(T_0, v_\nu)$, one could directly minimize $F(T, V, u_1, n_1)$.
with the condition of fixed $V = n_l v_l + n_v v_v$ 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

$$g_l(T_0, p) = g_v(T_0, p), \quad n_l v_l(T_0, p) + n_v v_v(T_0, p) = V,$$

which together determine the equilibrium pressure $p$ and the numbers $n_l$ and $n_v$ of moles in the two phases.
LECTURE VIII (TMD)

Fundamental relations in the representations of $F$, $G$ and $H$ functions and of the Massieu-Planck 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 a suitable reservoirs representing the surrounding), its equilibrium states are determined by minima (over the set of all possible virtual states) of the Helmholtz 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 roles of thermodynamic potentials. We have also argued that if 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 $n_1, \ldots, n_r$), the thermodynamical information about the system if complete. We now want to argue that if the thermodynamic potentials are known as functions of their natural variables: $F$ as a function of $T$, $V$ and $n$ (or $n_1, \ldots, n_r$), $H$ as a function of $S$, $p$ and $n$ or $G$ as a function of $T$, $p$ and $n$, the thermodynamical 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 controled (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 which is precisely the way of switching from a given function $f(x)$ to another function $g(p)$, where $p$ 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\footnote{The generalization of the Legendre transform to nonconvex functions is called the Fenchel transform and is defined as $g(p) = \sup_x (f(x) - px)$ and plays the important role in the theory of phase transitions.} functions (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

$$\tilde{f}(p) \equiv f(x(p)).$$ \hfill (135)

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
expressing it in terms of \( S \). But switching to \( \tilde{f}(p) \) one loses 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 9).

However the set of lines tangent at every point to a given a 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 10):

\[
y = p x + f(x_0) - p x_0 .
\]

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 (135) we will get the function

\[
g(p) = f(x(p)) - p x(p) ,
\]

which contains the same information about the dependence of the quantity \( f \) on the variable \( x \) as the original function \( f(x) \), but is expressed through the variable which is easily controlled experimentally. That the combination \( f(x) - px \) is indeed a function of the variable \( p \) can be also seen by considering its differential

\[
dg \equiv d(f(x) - px) = df - d(px) = \frac{df}{dx} dx - p dx - xd p = -x dp .
\]

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.

Of course a function of several variables, \( f = f(x_1, \ldots, x_n) \) can be Legendre transformed in an arbitrary number of its variables to obtain, say,

\[
g(p_1, \ldots, p_k, x_{k+1}, \ldots, x_n) = f(x_1, \ldots, x_n) - \sum_{j=1}^{k} p_j x_j ,
\]

Figure 9: The family of functions \( f_a(x) \) which all yield the same function \( \tilde{f}(x(p)) \) where \( p \) is the slope of the tangent line.
where the variables \(x_1, \ldots, x_k\) should be expressed in terms of the variables \(p_i\) by inverting with respect to them the \(k\) relations

\[
p_i = \frac{\partial f}{\partial x_i}.
\]

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, i.e. to trade the generalized variables \(q^l\) and \(\dot{q}^l\), \(l = 1, \ldots, n\) for the variables \(q^l\) and \(p_l = \partial L/\partial \dot{q}^l\). (Because of this, students have the tendency to think that the Legendre\(^\text{trans}^{65}\) transform has something to do with the fact that in mechanics variables come in pairs \(q^l\) and \(\dot{q}^l\). The derivation above clearly shows this is not so.) The Hamiltonian \(H(p_1, \ldots, p_n, q^1, \ldots, q^n)\) is just the (minus, in order that it has - in most cases the interpretation of mechanical energy) Legendre transform of the Lagrangian function \(L = L(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)\) in \(n\) its last variables

\[
H(p_1, \ldots, p_n, q^1, \ldots, q^n) = \sum_{j=1}^{n} p_j \dot{q}^j - L(q^1, \ldots, q^n, q^1, \ldots, q^n).
\]

It is also interesting to illustrate the working of the Legendre transform on another physical example taken from electrostatics. Let us determine the force by which the two plates of a capacitor attract each other. We remember that the voltage \(\varphi\) between the plates is related to the charge \(Q\) and the capacity \(C\) by \(Q = C \varphi\). If the capacitor is being chaged by successively bringing onto its plate infinitesimal portions \(dQ\) of charge (e.g. in the flat capacitor by transporting successively and reversibly - in the sense that an external force only counteracts the electric forces acting on the transported charge - portions \(dQ\) from one plate to the other) the work \(dW\) done on it by external forces is equal \(dW = \varphi\ dQ\). Charging in this way the capacitor with the charge \(Q\) requires doing

\[^{65}\text{Legendre, Lagrange, Laplace, Lavoisier - who can distinguish all these French L-masters?}\]
on it the work

\[ \int_0^Q dW = \int_0^Q \varphi dQ = \int_0^Q \frac{Q}{\mathcal{C}} dQ = \frac{Q^2}{2\mathcal{C}} \equiv U, \]

which is therefore equal to the energy of the charged capacitor (its internal energy if we
treat the capacitor in the thermodynamic way).

If the capacity \( \mathcal{C} \) of the capacitor is altered as a result of the action of an external
force \( F \) (when the application of the external force makes possible changing 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 of the capacitor, provided the
capacitor is an isolated system, that is, has fixed charge \( Q \) (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 virtual works principle but in
fact this is precisely the same as applying 1TMDL to isolated systems - in thermodynamics
systems adiathermally isolated; here, however, 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 constancy of \( Q \).) Let us write this statement in the thermodynamical
way

\[ dU(Q, \mathcal{C}) = \left( \frac{\partial U}{\partial Q} \right)_\varphi dQ + \left( \frac{\partial U}{\partial \mathcal{C}} \right)_Q d\mathcal{C} \equiv \varphi dQ + \mathcal{F} d\mathcal{C}, \]

By \( \mathcal{F} \) we have denoted here the generalized force related to the change of the capacity at
fixed charge on the capacitor. Expressing the capacity \( \mathcal{C} \) through the spatial characteris-
tics of the capacitor (the area of its plates, the distance between them) allows to give the
gearlized force \( \mathcal{F} \) the ordinary mechanical meaning. For example if \( \mathcal{C} \) changes due to
changing the distance between the plates

\[ dU(Q, \mathcal{C})|_{Q=\text{const.}} = \mathcal{F} d\mathcal{C} = \mathcal{F} \frac{d\mathcal{C}}{dz} dz \equiv F_z dz. \]

In this way we find that the external force \( F_z \) which precisely balances the force 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 \( \mathcal{C} = \varepsilon_0 (A/z) \)) attract
each other is given by

\[ F_z = \left( \frac{\partial U}{\partial \mathcal{C}} \right)_Q \frac{d\mathcal{C}}{dz} = \left( \frac{\partial U}{\partial z} \right)_Q = \frac{Q^2}{2\varepsilon_0 A}. \]

The force \( F_z \) found in this way cannot depend, of course, on whether the capacitor
has fixed charge \( Q \), or whether it is connected to a battery which keeps it at the voltage
\( \varphi \) of such a magnitude that the charge on the plates is equal \( Q \). If we write, however, the
energy \( U \) as a function \( U(\varphi, \mathcal{C}) \), expressing simply the charge through the voltage and the
capacity and try to define the generalized force $\mathcal{F}$ acting between the capacitor plates at constant voltage $\varphi$ by the relation

$$dU(\varphi, \mathcal{C})|_{\varphi=\text{const.}} = \left( \frac{\partial U}{\partial \mathcal{C}} \right)_\varphi d\mathcal{C} \equiv \mathcal{F} d\mathcal{C} \quad \text{(incorrectly)},$$

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

$$\tilde{U}(\varphi, \mathcal{C}) = U(Q(\varphi, \mathcal{C}), \mathcal{C}) - \varphi Q(\varphi, \mathcal{C}),$$

whose natural differential is

$$d\tilde{U}(\varphi, \mathcal{C}) = \left( \frac{\partial \tilde{U}}{\partial \varphi} \right)_\mathcal{C} d\varphi + \left( \frac{\partial \tilde{U}}{\partial \mathcal{C}} \right)_\varphi d\mathcal{C} \equiv -Q d\varphi + \tilde{\mathcal{F}} d\mathcal{C},$$

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 $\varphi$

$$d\tilde{U}(\varphi, \mathcal{C})|_{\varphi=\text{const.}} = \left( \frac{\partial \tilde{U}}{\partial \mathcal{C}} \right) \left[ \frac{1}{2} \mathcal{C} \varphi^2 - \varphi Q(\varphi, \mathcal{C}) \right] d\mathcal{C}$$

$$= \left( \frac{\partial \tilde{U}}{\partial \mathcal{C}} \right) \left[ -\frac{1}{2} \mathcal{C} \varphi^2 \right] d\mathcal{C} = -\frac{1}{2} \varphi^2 \frac{d\mathcal{C}}{dz} dz \equiv F_z dz.$$

In this way one gets the right expression for $F_z$ (with the right sign).

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 $\mathcal{C}$ 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

$$F_z dz + dW_{\text{bat}} = dU(\varphi, \mathcal{C}) = \frac{1}{2} \varphi^2 \frac{d\mathcal{C}}{dz} dz.$$

The work done by the battery consists of supplying the additional charge $\varphi d\mathcal{C}$ to the capacitor at the voltage $\varphi$; therefore this work equals $dW_{\text{bat}} = \varphi^2 d\mathcal{C}$. 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 $\varphi$. 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 with a heat bath at fixed temperature $T$.

In thermodynamics two families of potentials can be constructed. The first one (more commonly used) is obtained starting from the fundamental relation (98) in the internal energy representation $U = U(S, V, \ldots, n_1, \ldots, n_r)$. In the case of a simple system the

98
three basic potentials are the already introduced enthalpy (to simplify the notation \( n \) stands for \( n_1, \ldots, n_r \) and \( \mu \, dn \) for \( \sum_{j=1}^{r} \mu_j d\eta_j \))

\[
H(S, p, n) = U + pV, \\
dH = T(S, p, n) dS + V(S, p, n) dp + \mu(S, p, n) dn,
\]

the Helmholtz free energy

\[
F(T, V, n) = U - TS, \\
dF = -S(T, V, n) dT - p(T, V, n) dV + \mu(T, V, n) dn,
\]

and the Gibbs function (the free enthalpy)

\[
G(T, p, n) = U - TS + pV, \\
dG = -S(T, p, n) dT + V(T, p, n) dp + \mu(T, p, n) dn.
\]

One also defines the Grand potential

\[
\Omega(T, V, \mu) = U - TS - \mu n, \\
d\Omega = -S(T, V, \mu) dT - p(T, V, \mu) dV - n(T, V, \mu) d\mu.
\]

In the similar way one defines the functions \( H, F \) and \( G \) pertaining to other simple systems like wires and rubber bands \((-p \to K, V \to L)\) films \((-p \to \gamma V \to A)\) of magnetics
\((-p \to H_0 \) in the normal Gauss system or \(-p \to \mu_0 H_0 \) in the SI system, but if \( \mu \) 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 whose internal energy depends, in addition to the number of moles, of \( S, V \) and \( M \) one can form more Legendre transforms and their names are not codified. We have already used one of such potentials 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 full 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 presentations of enthalpy, free energy and Gibbs, respectively. To show this it is sufficient to realize that from each of them \( U = U(S, V, \ldots, n) \) 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

\[
U(T, V, n) = F + TS = F - T \left( \frac{\partial F}{\partial T} \right)_{V,n} \equiv -T^2 \left( \frac{\partial}{\partial T} \frac{F}{T} \right)_{V,n},
\]

and inverting the relation \( S = -(\partial F/\partial T)_{V,n} \) to get \( T = T(S, V, n) \) obtain \( U = (S, V, n) \). Moreover the potentials allow 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:

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

The distinguished role of the potentials $F$ and $\Omega$ stems from the fact that equilibrium statistical physics gives direct prescriptions to determine them on the basis of the microscopic dynamics (classical or quantum) of the considered system subject to specific constraints: the free $F$ Helmholtz energy if the system is isolated from the external pressure but remains in thermal contact with a heat bath at temperature $T$ and the potential $\Omega$ in the case of systems exchanging matter with a reservoir at the chemical potential $\mu$.

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 (93) whose differential

$$dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} dn \equiv \theta dU + \eta dV - \nu dn , \quad (147)$$

defines the variables $\theta$, $\eta$ and $\nu$ 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 them here any more.

Returning to the usual potentials it is good to notice that although the mechanical quantities like $p$ in the differentials $dU$ and $dF$ are denoted by the same letter, they are in principle functions of different variables (the natural ones for the respective potentials) and their experimental determination requires specifying the thermodynamical conditions. As an example consider the wire stretched by the force $K$ and satisfying the Hooke’s law. If we are given its internal energy $U = U(S, L)$, the tension $K$ is defined as the force needed to reversibly stretch the wire under adiathermal conditions, that is at constant entropy of the wire: $dU|_{S=\text{const}} = K dL$. In this way $K = K(S, L)$. The coefficient $k$ in the Hooke’s law which in mechanics is translated into the formula for the wire potential energy $E_{\text{pot}} = \frac{1}{2}k(L - L_0)^2$ is however measured usually by stretching the wire at constant temperature $T$ (of the environment). In such conditions the (minimal) work which must be done by RWS to stretch the wire from its equilibrium state with the length $L_0$ is, as discussed in the preceding Lecture, given by $\Delta F$ at fixed $T$. in this way measured is the coefficient $k(T)$ in the differential (written using the Hooke’s law)

$$dF = -SdT + k(T)(L - L_0)dL .$$

Integrating this formula one gets

$$F(T, L) = F(T, L_0) + \frac{1}{2} k(T)(L - L_0)^2 .$$

100
Since $S = -(\partial F/\partial T)_L$, one obtains

$$S(T, L) = -\frac{dF(T, L_0)}{dT} - \frac{1}{2} \frac{dK}{dT} (L - L_0)^2 \equiv S(T, L_0) - \frac{1}{2} \frac{dK}{dT} (L - L_0)^2.$$ 

Applying now the formula (146) one obtains

$$U(T, L) = F(T, L_0) + TS(T, L_0) + \frac{1}{2} \left( k(T) - T \frac{dK(T)}{dT} \right) (L - L_0)^2$$

$$= U(T, L_0) + \frac{1}{2} \left( k(T) - T \frac{dK(T)}{dT} \right) (L - L_0)^2.$$ 

Therefore the factor $k$ defined as the coefficient of $\frac{1}{2}(L - L_0)^2$ in the internal energy of the wire (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 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). Together they provide sufficient information on the coefficients of the differentials in the forms $dU = C_VdT + (\partial U/\partial V)_T dV$ (or $dU = (\partial U/\partial T)_p dT + (\partial U/\partial p)_{T,p} dp$) and $dS = (C_V/T) dT + (\partial S/\partial V)_T dV$ (or $dS = (\partial S/\partial T)_p dT + (\partial S/\partial p)_{T,p} dp$) 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 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 $dW = \mathcal{H}_0 dM$ can be (reversibly) done on the system in addition to the usual volume work $-p dV$, would require determining from data one more function. Similarly, reconstructing the dependence on the number(s) of moles $n$ ($n_i, i = 1, \ldots, r$) would in principle require supplemenary 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) puts, 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 the 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 derive the relations which follow from extensiveness of the system.

---

66All such functions are called the equations of state. Similarly as e.g. $C_V$, whose dependence on the volume 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 $U$ and $S$ are state functions) which, however, do not determine them completely, leaving room for experimental input.
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:

\[ S(\lambda U, \lambda V, \ldots, \lambda n_1, \ldots, \lambda n_r) = \lambda S(U, V, \ldots, n_1, \ldots, n_r). \]  

(148)

The same relation applies of course to \( U = U(S, V, \ldots, n_1, \ldots, n_r) \). One of the consequences, already discussed in Lecture VI, is that

\[ S = n s(u, v, \ldots, x_1, \ldots, x_r), \quad U = n u(s, v, \ldots, x_1, \ldots, x_r), \]

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 molar fractions. It also follows that the intensive parameters \( T, p, \mu_i \) which are defined as derivatives of \( U \) with respect to the extensive parameters must be homogeneous functions of order zero:

\[ T = T(\lambda S, \lambda V, \ldots, \lambda n_1, \ldots) = T(S, V, \ldots, n_1, \ldots) = T(s, v, \ldots, x_1, \ldots), \]

\[ p = p(\lambda S, \lambda V, \ldots, \lambda n_1, \ldots) = p(S, V, \ldots, n_1, \ldots, n_r) = p(s, v, \ldots, x_1, \ldots), \]

\[ \mu_i = \mu_i(\lambda S, \lambda V, \ldots, \lambda n_1, \ldots) = \mu_i(S, V, \ldots, n_1, \ldots, n_r) = \mu_i(s, v, \ldots, x_1, \ldots). \]  

(149)

Since \( x_1 + \ldots + x_r = 1 \), the \( 1 + o + r \) intensive parameters (here \( o \) is the number of works which can be done on the system) depends on only \( 1 + o + r - 1 = o + r \) variables, which implies that there must be one relation linking the intensive variables \( T, p, \ldots, \mu_1, \ldots, \mu_r \). If there is only one component (\( r = 1 \)), this relation uniquely determines the single chemical potential \( \mu \) as a function of the \( 1 + o \) intensive parameters \( T, p, \ldots \).

Furthermore, differentiating the relation (148) and the analogous relation written for \( U \) with respect to \( \lambda \) and setting then \( \lambda = 1 \), one obtains that

\[ U(S, V, n_1, \ldots) = ST(S, V, n_1, \ldots) - V p(S, V, n_1, \ldots) + \sum_{j=1}^{r} n_j \mu_j(S, V, n_1, \ldots), \]  

(150)

\[ S(U, V, n_1, \ldots) = U\theta(U, V, n_1, \ldots) + V \eta(S, V, n_1, \ldots) - \sum_{j=1}^{r} n_j \nu_j(S, V, n_1, \ldots), \]

(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 \( U = U(S, V, n_1, \ldots) \) and then using the form (150) of \( U \) we get

\[ dU = TdS - p dV + \sum_{j=1}^{r} \mu_j dn_j, \]

\[ dU = TdT + S dT - p dV - V dp + \sum_{j=1}^{r} \mu_j dn_j + \sum_{j=1}^{r} n_j d\mu_j. \]

Subtracting now these two expressions side by side leads to the formula

\[ S dT - V dp + \sum_{j=1}^{r} n_j d\mu_j = 0, \]  

(151)

102
known as the **Gibbs-Duhem relation.** Analogous operations done on the entropy written in two different ways give

\[ Ud\left(\frac{1}{T}\right) + Vd\left(\frac{p}{T}\right) - \sum_{j=1}^{r} n_j d\left(\frac{\mu_j}{T}\right) = 0. \]  

(152)

Dividing (151) by the total number \( n \) of moles gives the correlation of the differentials of the intensive parameters which depends only on molar quantities:

\[ s dT - v dp + \sum_{j=1}^{r} x_j d\mu_j = 0, \]  

(153)

and expresses the already noted fact that of \( 1 + o + r \) intensive parameters characterizing an extensive system, only \( o + r \) 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 (153) in written in the form

\[ d\mu = -s(T, p) dT + v(T, p) dp, \]  

(154)

can be 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 (150) of the internal energy it immediately follows that

\[ H(S, p, n_1, \ldots) = ST(S, p, n_1, \ldots) + \sum_{j=1}^{r} n_j \mu_j(S, p, n_1, \ldots), \]

\[ F(T, V, n_1, \ldots) = -Vp(T, V, n_1, \ldots) + \sum_{j=1}^{r} n_j \mu_j(T, V, n_1, \ldots), \]

\[ G(T, p, n_1, \ldots) = \sum_{j=1}^{r} n_j \mu_j(T, p, n_1, \ldots), \]  

(155)

\[ \Omega(T, V, \mu_1, \ldots) = -Vp(T, V, \mu_1, \ldots), \]

The same relations follow, of course, from the scaling properties of these functions: e.g. differentiating with respect to \( \lambda \) the relation

\[ F(T, \lambda V, \lambda n_1, \ldots) = \lambda F(T, V, n_1, \ldots), \]

and setting \( \lambda = 1 \) the second relation (155) is obtained. If there is only one component in the system, the third of these relations, after division by \( n \) gives

\[ \mu(T, p) = g(T, p) \equiv \frac{1}{n} G(T, p, n) = u - T s + p v. \]  

(156)

103
As advertised in the preceding Lecture, in this case the chemical potential is just the molar Gibbs function. The formula (156) gives another method of calculating the chemical potential of a system composed of one component only.

If the system is composed of more than one component, the Gibbs function $G$ is given by (155) as the sum of chemical potentials weighted by the respective mole numbers, or

$$g(T, p, x_1, \ldots, x_r) = \sum_{j=1}^{r} x_j \mu_j(T, p, x_1, \ldots, x_r), \quad (157)$$

and determination of the individual chemical potentials requires more experimental input.

**Mixture of perfect gases and 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 of such a mixture in equilibrium is, when **expressed through temperature and 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):

$$U(T, n_1, \ldots, n_r) = \sum_{i=1}^{r} U_i(T, n_i) = \sum_{i=1}^{r} n_i u_i(T) = n \sum_{i=1}^{r} x_i u_i(T),$$

$$S(T, V, n_1, \ldots, n_r) = \sum_{i=1}^{r} S_i(T, V, n_i) = \sum_{i=1}^{r} n_i s_i(T, v_i) = n \sum_{i=1}^{r} x_i s_i(T, v_i), \quad (158)$$

where $n = n_1 + \ldots + n_r$, $x_i = n_i/n$ and $v_i = V/n_r$.

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 (93): 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 formula $dS = (dU + p \, dV)/T$ is also satisfied:

$$dS = \sum_i \frac{\partial S_i}{\partial U_i} dU_i + \sum_i \frac{\partial S_i}{\partial V} dV = \sum_i \frac{1}{T_i} dU_i + \sum_i \frac{p_i}{T_i} dV$$

$$= \frac{1}{T} d\left( \sum_i U_i \right) + \frac{1}{T} \left( \sum_i p_i \right) dV = \frac{1}{T} dU + \frac{p}{T} dV.$$
The immediate consequences of the postulate are the formulae for the molar heat capacity of the mixture:

\[ c_v^{\text{mix}} = \sum_{i=1}^{r} x_i c_v^{(i)}, \quad \text{bo} \quad c_v^{\text{mix}} \equiv \frac{1}{n} \left( \frac{\partial U}{\partial T} \right)_V, \]

and its pressure (we remember the formula for \( S(T, V, n) \) of the perfect gas!):

\[ p^{\text{mix}} = \sum_{i=1}^{r} p_i \quad \text{bo} \quad \frac{p^{\text{mix}}}{T} \equiv \left( \frac{\partial S}{\partial V} \right)_{U,n_i} = \sum_{i=1}^{r} n_i \frac{R}{V}, \]

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 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 verified: the first one by trivial measurement of the molar heat capacities of individual gases and of its mixture and the second one by the device called the van ’t Hoff box shown in Figure 11 - 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 whose mutual interactions are negligibly weak\(^{67}\) is solvable within the so called

\(^{67}\)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 neutrinos have indeed different temperatures.

105
canonical ensemble approach which gives the Helmholtz function \( F(T, V) \) of the system: in the case of nointeracting gases it automatically gives \( F \) of the form

\[
F(T, V, n_1, \ldots, n_r) = \sum_{i=1}^{r} F_i(T, V, n_i).
\]

We know (we already should!), that the function \( F(T, V, n_1, \ldots, n_r) \) contains information about all termodynamical characteristics of the system.

One more consequence of the Gibbsa postulate is the **entropy of mixing**, which is easily identified when the formula (158):

\[
S(T, V, n_1, \ldots, n_r) = \sum_{i=1}^{r} n_i \left( s_i(T_0, v_0) + \int_{T_0}^{T} \frac{dT'}{T'} c_{v_i}^{(i)}(T') + R \ln \frac{v_i}{v_0} \right),
\]

for the entropy of the mixture (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 \( n \):

\[
S(T, V, n, x_1, \ldots, x_r) = n s^{\text{mix}}(T_0, v_0) + n \int_{T_0}^{T} \frac{dT'}{T'} c_v^{\text{mix}}(T') + nR \ln \frac{V}{nv_0}
\]

\[+ n \sum_{i=1}^{r} (-R x_i \ln x_i). \tag{159}\]

(We have introduced here \( s^{\text{mix}}(T_0, v_0) = \sum_i x_i s_i(T_0, v_0). \) The last term is the mixing entropy which we will also call the Cinderella’s (Kopciuszek) entropy. It is positive because \( x_i \leq 1 \). That this term is related to the mixing of different gases in one container follows clearly from the comparison of (159) with the entropy of \( n \) moles of a one-component perfect gas which (accidentally) would have the 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). A one which offers an insight into its origin is as follows. Prepare \( n_1 \) and \( n_2 \) moles of two different 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

\[
\mu_i(T, p, x_1, \ldots, x_r) = \mu(T, x_ip) \equiv \mu(T, p_i) = \mu(T, p) + RT \ln x_i, \quad (160)
\]

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 chemical potentials of the two components can be obtained without additional experimental (or statistical physics) input is the 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 of \( T, p \) and \( x \), can be then expanded in the Taylor series:

\[
u(T, p, x) = u_0(T, p) + x \Delta u(T, p) + \ldots,
\]
\[
v(T, p, x) = v_0(T, p) + x \Delta v(T, p) + \ldots,
\]
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). Moreover, the first bracket must also be an exact differential of the molar entropy of the pure solvent. It follows, that also the second bracket must have this property and the one-form \( ds \) can be integrated up 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 temperature 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:

\[
g(T, p, x) = u - Ts + pv = g_0(T, p) + x \Delta g(T, p) + RT \{(1 - x) \ln(1 - x) + x \ln x\},
\]
in which \( g_0 = u_0 - T s_0 + p v_0 \) 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 \( (nx = n_2, n(1 - x) = n_1) \)

\[
G(T, p, n_1, n_2) = (n_1 + n_2) g(T, p, x) = (n_1 + n_2) g_0(T, p) + n_2 \Delta g(T, p) + RT n_1 \ln \frac{n_1}{n_1 + n_2} + RT n_2 \ln \frac{n_2}{n_1 + n_2}.
\]

The chemical potentials can be then obtained as the derivatives

\[
\mu_1 = \left( \frac{\partial G}{\partial n_1} \right)_{T, p, n_2} = g_0(T, p) + RT \ln(1 - x) \approx g_0(T, p) - RT x,
\]

\[
\mu_2 = \left( \frac{\partial G}{\partial n_2} \right)_{T, p, n_1} = g_0(T, p) + \Delta g(T, p) + RT \ln x \equiv \psi(T, p) + RT \ln x. \quad (161)
\]

108
LECTURE IX (TMD)

Phase transitions
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. In some range of intensive parameters a given phase is stable in another it may become metastable and in yet some other ranges it simply cannot exist. If in certain conditions characterized by the values of temperature and pressure (and of other intensive parameters, if the substance is not simple) there may exist as stable more than one phase, one speaks of the coexistence of phases. It is in these conditions that phase transitions normally occur.\textsuperscript{68} Phases of a given substance should not be confused with its physical states (solid, liquid and gaseous): a given solid can have different phases (called allotropic modifications) in which atoms are differently ordered in the crystalline lattice cells: for example, silicon Si has 3 different allotropic phases, tin Sn two (white tetragonal and grey which is amorphous), H\textsubscript{2}O has as much as 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.

If in given conditions a substance may exist in more than one phase, the system is usually inhomogeneous - the phases exist in spatially separated forms. They are then treated as different parts of a compound system separated by (fictitious) walls allowing for transfer of matter, energy and volume between the parts.\textsuperscript{69}

The conditions in which 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 $\alpha$ and $\beta$ of such a substance always required the equality of the temperatures (thermal equilibrium with respect to exchanging energy), pressures (mechanical equilibrium with respect to exchanging volume) and the equality of the chemical potentials of the two phases. This extends to multicomponent systems, as well, the only change being that in such cases the chemical potentials characterizing the same component of both phases must be pairwise equal:

$$\mu_i^{(\alpha)}(T, p, x_{1}^{(\alpha)}, \ldots, x_{r-1}^{(\alpha)}) = \mu_i^{(\beta)}(T, p, x_{1}^{(\beta)}, \ldots, x_{r-1}^{(\beta)}), \quad i = 1, \ldots, r.$$  \hspace{1cm} (162)

The difference between open and isolated systems is that in the former case the equilibrium of phases can be neutral (the first example considered in Lecture VIII), while in the latter case it can uniquely determine the partition of matter components between the phases.

\textsuperscript{68}A phase which in a given range of the parameters is metastable may exist for a very 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.

\textsuperscript{69}The boundary separating the phases, although neglected in further discussion here, has in fact some thickness; it can be treated as consisting of some quantity 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 be also modified in presence of external fields like the gravitational one.
The questions how many phases of a given substance can simultaneously be in equilibrium finds the answer in the Gibbs phase rule. Let \( r \) be the number of components (labeled by the index \( i \)) of a simple substance which can exist in \( f \) phases (labeled by the superscript \( \alpha \)). If each phase is an isolated system, it is characterized by \( 2 + r \) extensive parameters: its entropy \( S^{(\alpha)} \), its volume \( V^{(\alpha)} \) and the numbers of moles \( n_1^{(\alpha)}, \ldots, n_r^{(\alpha)} \) 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. Firstly, they must have all the same temperatures (thermal equilibrium)

\[
T^{(1)} = T^{(2)} = \ldots = T^{(f)} \equiv T,
\]

which gives \( f - 1 \) conditions. Another \( f - 1 \) conditions follow from the equality of their pressures (mechanical equilibrium):

\[
p^{(1)} = p^{(2)} = \ldots = p^{(f)} \equiv p.
\]

Equalities of the chemical potentials (chemical equilibrium with respect to matter exchange)

\[
\mu_1^{(1)} = \mu_1^{(2)} = \ldots = \mu_1^{(f)} \equiv \mu_1,
\]

\[
\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \l
Thus two phases \((f = 2)\) of a single component \((r = 1)\) substance (like pure \(\text{H}_2\text{O}\)) 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. If chemical reactions are allowed to occur between the components of the system, each possible reaction imposes one condition on the molar fractions and the number of degrees of freedom decreases by one per each reaction.

Below we will discuss only phase transitions of one-component simple substances. Suppose a quantity \((n\) moles) of such a substance is in a state \((\text{characterized by its internal temperature } T \text{ and pressure } p)\) 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 \(p\) assume values at which two (or three) phases can exist, the proportion of the substance in the phases will change - the phase transition will occur.

**Classification of phase transitions**

Two phases can coexist only if their chemical potentials are equal. Thus the chemical potentials are continuous along the coexistence lines. P. Ehrenfest classified phase transitions according to the continuity of derivatives of chemical potentials: In first order transitions continuous are chemical potentials but their first derivatives (at least one of them) - the molar entropies and/or molar volumes are discontinuous. In second order phase transitions the chemical potentials and their first derivatives, \(s\) and \(v\), are continuous but at least one of the derivatives of \(s\) or \(v\), that is, \(c_p\), \(\alpha_p\) or \(k_T\), is discontinuous, and so on.

Of the transitions which fit 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 \(\mathcal{H}\). 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\(^7^1\) phase diagram of a simple, single component substance which when conditions are varied, undergoes first-order phase transitions (perhaps with the exception of an isolated point(s)). The diagram shows regions on the \((T, p)\) plane of stable phases. Phase diagrams of multicomponent substances can be considerably more complicated: apart from \(T\) and \(p\) also molar fractions \(x_i\) enter the game and one can only display projections of stability regions onto various two-dimensional planes.

We consider a simple 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

---

\(^7^1\)Atypical phase diagrams will be discussed later.
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 12 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 (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 variation of these functions with \( T \) and \( p \) which readily follow from the formulae

\[
dg \equiv d\mu = -s \, dT + v \, dp, \quad \left( \frac{\partial g}{\partial T} \right)_p = -s, \quad \left( \frac{\partial g}{\partial p} \right)_T = v.
\]

Thus in the direction of increasing pressure the \( g \)-surfaces (\( \mu \)-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\(^{72} \) 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 (whose Gibbs function falls most steeply as the pressure goes down), but as far as the

\(^{72}\text{Of course, for each substance there is only one arbitrary constant entropy } s_0.\)
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 pressure - closer to the $p$ axis than the solid phase). In the phase diagram shown in Figure 12a crossing the line $tr\text{-}S$ 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 $v^{(\text{solid})} < v^{(\text{liquid})}$; the substance expands on melting. Figure 12b shows the opposite situation in which $v^{(\text{solid})} > v^{(\text{liquid})}$: here the substance (water is here the most obvious example) expands on solidifying. When a coexistence line in Figure 12a or 12b 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 12 the transitions solid$\leftrightarrow$liquid, solid$\leftrightarrow$vapour or liquid$\leftrightarrow$vapour heat is absorbed. The finite difference of entropies of the 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 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 vertical 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 consider 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 $-40^\circ\text{C}$. And conversely, liquids can be superheated - another historic device of particle physics, the bubble chamber exploited this possibility. 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. 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 12 - 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).

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 also be 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 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

$$\frac{dp}{dT} = \frac{s^{(\beta)} - s^{(\alpha)}}{v^{(\beta)} - v^{(\alpha)}} = \frac{q_{\alpha \to \beta}}{T(v^{(\beta)} - v^{(\alpha)})},$$

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: e.g. 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 $\rho = m_{\text{mol}}/v$) of the solid and liquid phases. Water molar volume is smaller than that of ice (which has also lower entropy - the latent heat $q_{\alpha \to \beta}$ is defined to be positive, so $\alpha = \text{ice}, \beta = \text{water}$) which results in the negative slope (negative right hand side of the equation (164)) of the ice melting curve (which is of the type shown in Figure 12b). 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) was understood by the second Thomson, the brother of the lord Kelvin, around the middle of the XIX century to be required by 2TMDL. If it were vertical, one could devise an “ice-engine”, shown in figure ??, which would violate the Kelvin’s 2TMDL: starting with a quantity of water kept at $0^\circ \text{C}$ one could place a weight $m$ on the piston closing the vessel with the water from above (the water compressibility is small and can be neglected here). Then a quantity $\dot{Q}$ could be extracted from the water to a heat reservoir at $0^\circ \text{C}$, 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 $0^\circ \text{C}$) 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 \( W = mgh \) would be done at the cost of the heat \( Q - \bar{Q} \) taken entirely from the single reservoir at \( 0^\circ \text{C} \). 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 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 Clapeyron equation was one of the first successful applications of thermodynamics to physical problems and lend a great confidence to this developing branch of theoretical physics.

**Critical point**

The line of coexistence of the liquid and vapour phases (the line \( \text{tr-cr} \) on the diagrams 12a,b) does not continue indefinitely: as the line is followed to the right, the quantitative characteristics (densities, heat capacities, compressibilities, etc.) of the vapour and liquid phases become more and more similar to eventually disappear at the critical point marked “cr” (and characterized by \( T_{\text{cr}} \) and \( p_{\text{cr}} \)) where the liquid and vapour become indistinguishable. 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 \( T < T_{\text{cr}} \) is compressed isothermally (one moves vertically upwards on the phase diagrams 12a,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 the volume does not increase pressure (\( k_T \) stays infinite), only increases the quantity of liquid as compared to the quantity of the vapour (on the diagram 12 one is staying all this time on the coexistence curve); when the vapour has dissapeared completely, the pressure begins to rise rather steeply (\( k_T \) falls abruptly to a small value) and one moves on the diagram 12 upwards already above the coexistence line.

When the analogous compression of the vapour is effected exactly at the critical temperature \( T_{\text{cr}} \), the compressibility \( k_T \) first decreases as the volume is reduced (and pressure rises) then rises to become momentarily infinite at \( p_{\text{cr}} \) and then steadily goes down. At no point is there a separation of the system into two phases is observed. Above \( T_{\text{cr}} \) 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 12 to make the system to pass from the vapour phase to the liquid phase, without any discontinuous changes in its properties.\(^{73}\) This means that the two \( g \)-surfaces: the \( g_1(T,p) \)

\[^{73}\text{It is appropriate to make a comment on the terminology: in the older literature the system at temperatures } T < T_{\text{cr}} \text{ is called either liquid (at higher pressures) or vapour (at lower pressures) while the system above } T_{\text{cr}} \text{ is called gas. More adequate seems to call the system above } T_{\text{cr}} \text{ supercritical fluid and}\]
one and the \( g(T, p) \) one which we compared in establishing which of the phases is stable in the given domain of the \((T, p)\) plane, are in fact to parts of one and the same \( g \)-surface which simply intersects itself for \( T < T_{\text{cr}} \) along a line whose projection onto the \((T, p)\) plane is the curve tr-cr, but not for \( T > T_{\text{cr}} \).

One way of visualizing the form of such a self-intersecting \( g \)-surfaces is to use the Van der Waals equation of state

\[
(p + \frac{a}{v^2})(v - b) = RT,
\]

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 \( T > T_{\text{cr}} \) isotherms of a real vapour-liquid system. Below some temperature, which should therefore be identified with \( T_{\text{cr}} \), to a given pressure \( p \) there correspond three different values (of which the smallest one and the largest one we, introducing their interpretation, denote \( v_1 \) and \( v_\nu \), 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 \( T > T_{\text{cr}} \), but for \( T < T_{\text{cr}} \) it becomes negative somewhere between \( v_1 \) and \( v_\nu \) (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 those 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 12 from below (from above), the system breaks up into two phases and its number of moles \( n \) and the molar volume \( v \) are given by

\[
n = x_1 n_1 + x_\nu n_\nu, \quad v = x_1 v_1 + x_\nu v_\nu, \quad x_1 + x_\nu = 1,
\]

with \( x_\nu \) 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 12 from below);

the one below \( T_{\text{cr}} \) vapour or, excangeably, gas.
• the segment B-C can be interpreted as representing the supercooled vapour (a metastable state) - on Figure 12 the corresponding points lie above the tr-cr line;

• the segment C-D-E is unphysical ($kT < 0$);

• the segment E-F can be interpreted as representing the superheated liquid (a metastable state) - on Figure 12 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 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 $T < T_{cr}$ from the point, say A, on: $^{74}$

$$g(T, p) = g(T, p_A) + \int_{p_A}^{p} dp' \, v(T, p').$$  \hspace{1cm} (166)

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 $T = \text{const}$ 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 $T_{cr}$; 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 placed in such a way that the areas B-C-D-B and D-E-F-D should be equal. In this way the value of the molar Gibbs function at the point E (that is of the chemical potential of the vapour in equilibrium with its liquid at this value of $T$) will be the same as the value of the molar Gibbs function at the point F. Indeed, the Maxwell rule requires that (geometrically)

$$\text{(area under B - C)} - \text{(area under C - D)} = \text{(area under D - E)} - \text{(area under E - F)}$$

$^{74}$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.
But this precisely ensures that (recall how the integration (166) has been done)

\[
g(T, p_F) - g(T, p_B) = \int_{p_B}^{p_F} dp \nu(T, p) \\
= (\text{area under } B - C) - (\text{area under } C - D) \\
- (\text{area under } D - E) + (\text{area under } E - F) = 0. 
\]  

The construction is however of doubtful validity for it clearly exploits the unphysical part of the isotherm (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 (which does not involve the unphysical part of the isotherm) and at the same time to start form the point G and move back along the isotherm again computing at every point the chemical potential. In this way one can find the two points, one on the left part of the isotherm and the other on the its right part, at which the two potentials have the same value. It is 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 the thermodynamical functions (everywhere except for the 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.

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 (3.73 K, 0.14 K and 9 K, in the case of tin, iridium and niobium in zero magnetic field, respectively - we will not consider hight temperature superconductors here), that is exhibit zero electric resistance. This is accompanied, and one can show that the two effects must necessarily come together, 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. \( \mathbf{B} = \mathbf{0} \) inside the specimen; this is called the Meissner-Ochsenfeld effect) unless the strength of the external field is too strong: if it exceeds some critical value \( H_c \) (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 \( \mathbf{B} \) in the material is related to the strength \( H \) inside by some magnetic susceptibility \( \chi \): \( \mathbf{B} = H + 4\pi \mathcal{M} \) where \( \mathcal{M} = \chi H \). Below we will assume that the normal phase susceptibility \( \chi \) is negligibly small, so that in the normal state simply \( \mathbf{B} = H \) inside the material (just as outside it).

Suppose a tin specimen of a long cylindrical shape is placed in a weak external magnetic
field $^7$ at a temperature lower than the critical one, $T_c$ (so that the specimen is in its superconducting phase). The magnetic field induction inside the specimen is then zero. If $H$ is slowly increased so that it exceeds some critical value, the metal undergoes a first order (as will be seen) phase transition to the normal state in which $B = H$ inside it. If the external field is reduced back below the critical value $H_c$, 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 dependence of the critical field strength $H_c$ on the temperature (at fixed pressure) is conveniently modeled by the empirical formula

$$H_c(T) = H_0 \left[ 1 - \left( \frac{T}{T_c} \right)^2 \right], \quad (168)$$

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 $H = H_c$ (which is crucial for further considerations). The dependence on the pressure is rather weak: $T_c$ 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

$$g' = u' - Ts + pv - H M v, \quad (169)$$

(we assume that the the magnetic field strenght $H$ and the magnetization $M$ are constant in the material) whose differential $dg' = -sdT + vdp - vM dH$ implies that

$$\left( \frac{\partial g'}{\partial T} \right)_{p,H} = -s, \quad \left( \frac{\partial g'}{\partial p} \right)_{T,H} = v, \quad \left( \frac{\partial g'}{\partial H} \right)_{T,p} = -vM,$$

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 coexist along a two dimensional surface at which their molar gibbs functions are equal $g'_n = g'_s$. The coexistence surface projected onto the three planes: $(T, H)$, $(p, H)$ and the $(T, p)$ one, gives rise to three coexistence curves which are determined by the equations analogous to the Clapauius one - the method of derivation is exactly the same.$^8$

$$\left( \frac{\partial H_c}{\partial T} \right)_p = -\frac{s_n - s_s}{v_n M_n - v_s M_s},$$

$$\left( \frac{\partial H_c}{\partial p} \right)_T = \frac{v_n - v_s}{v_n M_n - v_s M_s}, \quad (170)$$

$$\left( \frac{\partial p}{\partial T} \right)_H = \frac{s_n - s_s}{v_n - v_s}.$$

---

$^7$Since we consider the magnetic field aligned always in the same direction, we can play with the scalar quantity $H$.

$^8$E.g. if the pressure is fixed, on the coexistence curve the changes $dH_c$ and $dT$ are correlated so that $g'_n(T + dT, H_c + dH_c, p) = g'_s(T + dT, H_c + dH_c, p)$. 

119
As in the superconducting (perfectly diamagnetic) phase \( B = 0 \) means that \( \mathcal{M}_s = -\mathcal{H} \), while the equality \( B = \mathcal{H} \) holding in the normal phase implies that in this phase \( \mathcal{M}_n = 0 \), the first two of the above three equations can be rewritten in the forms

\[
\left( \frac{\partial \mathcal{H}_c}{\partial T} \right)_p = -\frac{4\pi}{v_s \mathcal{H}_c} (s_n - s_s), \quad \left( \frac{\partial \mathcal{H}_c}{\partial p} \right)_T = -\frac{4\pi}{v_s \mathcal{H}_c} (v_n - v_s).
\] (171)

The values \( s_n, s_s, \) etc. should be taken on the coexistence curve, but their variation with the magnetic field is not large: Since, as readily follows from (169), \((\partial s/\partial H)_T,p = \mathcal{M} / (\partial \mathcal{M}(H) / \partial T)_p \) \( \mathcal{H} \), neither \( s_n \) nor \( s_s \) are appreciably field-dependent: the former because \( \mathcal{M}_n = 0 \) and the latter because \( \mathcal{M}_s = -\mathcal{H}/4\pi \), so only a rather small temperature dependence of \( v_s \) can contribute to the variation of \( s_s \) with the temperature. Similarly, since \((\partial v/\partial H)_T,p = -\mathcal{M} / (\partial \mathcal{M}(H) / \partial p)_T,H \), \( v_n \) is field independent and to the field dependence of \( v_s \) can contribute only the small compressibility of the specimen \( n \) this phase. Thus, with the exception of the difference \( v_n - v_s \) one can use in the above equations \( s_n \) and \( v_s \) taken at the zero magnetic field and similarly neglect the field dependence of \( s_s \) and \( v_s \).

Now, the first of the equations (171) rewritten in the form

\[
(s_n - s_s) = \frac{v_s \mathcal{H}_c}{4\pi} \left( \frac{\partial \mathcal{H}_c}{\partial T} \right)_p, \quad (v_n - v_s) = \frac{v_s \mathcal{H}_c}{4\pi} \left( \frac{\partial \mathcal{H}_c}{\partial p} \right)_T,
\] (172)

show, because as stressed and as the formula (168) makes it clear, the slope \((\partial \mathcal{H}_c/\partial T)_p \) is not infinite at \( \mathcal{H}_c = 0 \), that at \( \mathcal{H}_c = 0 \) the difference \( s_n - s_s \) vanishes. This shows that in zero magnetic field the transition conductor-superconductor occurs without any latent heat. In the same way the second equation shows that in zero magnetic field vanishes also the difference \( v_n - v_s \). Thus is in the ordinary two-parameter space \((T,p)\) corresponding to zero magnetic field, the conductor-superconductor transition cesses to be first order. This means that the third of the equations (170) is indeterminate at zero field.\(^7\) 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 (172) with respect to \( T \) at constant pressure (again neglecting the temperature dependence of \( v_s \) in their right hand sides) one finds

\[
(c_{p}^{(n)} - c_{p}^{(s)}) = \frac{v_s T}{8\pi} \left( \frac{\partial^2 \mathcal{H}_c^2}{\partial T^2} \right)_p = -\frac{v_s T}{4\pi} \left[ \mathcal{H}_c \left( \frac{\partial^2 \mathcal{H}_c}{\partial T^2} \right)_p + \left( \frac{\partial \mathcal{H}_c}{\partial T} \right)_p \right]^2.
\] (173)

Appealing to the empirical formula (168) one sees that the discontinuity of the molar heat capacities is finite at zero magnetic field and is equal

\[
(c_{p}^{(n)} - c_{p}^{(s)}) = -\frac{v_s \mathcal{H}_c^2}{\pi T_c}.
\]

\(^7\)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 (171) is well defined and does determine the coexistence curve.
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 (170) 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 (172) with respect to \(T\) and with respect to \(p\) (again neglecting any small dependence of \(v_s\) 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:

\[
\left(\frac{\partial H_c}{\partial T}\right)_p^2 = -\frac{4\pi}{\nu T} \left(c_p^{(n)} - c_p^{(s)}\right),
\]

\[
\left(\frac{\partial H_c}{\partial p}\right)_T^2 = -4\pi \left(k_T^{(n)} - k_T^{(s)}\right),
\]

\[
\left(\frac{\partial H_c}{\partial T}\right)_p \left(\frac{\partial H_c}{\partial p}\right)_T = 4\pi \left(\alpha_p^{(n)} - \alpha_p^{(s)}\right).
\]

Taking their ratios and exploiting the standard shocking relation 
\((\partial H_c/\partial p)_T(\partial H_c/\partial T)_p = -(\partial T_c/\partial p)_H\), one obtains the two (mutually consistent, as follows from the construction) equations

\[
\left(\frac{\partial T_c}{\partial p}\right)_{\mathcal{H}=0} = v_s T \left(\frac{\alpha_p^{(n)}}{c_p^{(n)} - c_p^{(s)}} - \frac{\alpha_p^{(s)}}{c_p^{(n)} - c_p^{(s)}}\right) = \frac{k_T^{(n)} - k_T^{(s)}}{\alpha_p^{(n)} - \alpha_p^{(s)}},
\]

which replace the last of the equations (170) 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.

\[78\text{Recall that } k_T = -(1/\nu)(\partial v/\partial p)_T \text{ and } \alpha_p = (1/\nu)(\partial v/\partial T)_p.\]
LECTURE X (STAT)

We have finished the thermodynamic part of this Course. Thermodynamics is a phenomenological theory whose main aim is establishing relations between measurable quantities characterizing macroscopic properties of bodies (systems). It also formulates general rules allowing to tell which processes are possible and imposes limits on possible 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) whose first (but not the only one!) goal is the derivation of thermodynamics or, put differently, providing methods for computing equilibrium properties of various physical systems on the basis of more fundamental laws79 (theories) governing the behaviour of microscopic constituents (molecules, atoms or physical fields) of these systems. 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 nevertheless does not prevent making definite predictions for macroscopic quantities characterizing a given system and to provide concrete recipes 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 possible to go further and relying on microscopic dynamical theories investigate also systems not in equilibrium, in particular 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.

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,

79 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.
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), the state of equilibrium. 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 (probably) the statement that at the microscopic level no real system can be absolutely isolated with respect to highly random (uncontrollable) minute external perturbations. This is particularly true of that macroscopic part of the system on which measurements are being made and which for short periods (when the measurements are made) can be treated as macroscopically isolated from the rest of the system. One can also adopt the view that, for reasons which will be given in the next Lecture, this is certainly true of any macroscopically large quantum system, so it is the right basis on which to base quantum statistical mechanics, while the classical one must be constructed in such a way to be the limit of the quantum one.

When we are interested (we measure) those few properties of a macroscopic system which characterize its equilibrium state, then the same their values can be obtained by measuring them on copies of the system which microscopically are in different states than our system. In other words, typically the same macrostate of the 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 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 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 assumption on which to base the foundations of statistical mechanics of a macroscopically isolated system is to assume that all possible microstates of such a system 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.\(^8\) For this reason one could think that the foundations of equilibrium and nonequilibrium statistical physics should be based only on this theory. While this is principally true, it is nevertheless worthwhile to consider statistical physics based on classical mechanics. Although some ingredients necessary to fully account for the behaviour of real physical systems are lost in this way (for example spin is lost, indistinguishability of identical particles can only approximately be accounted for, phenomena at very low,

---

\(^8\) Or, at a yet deeper level penetrated to date, by the laws of quantum field theory which is in a sense (it takes me roughly 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.
$T \to 0$, 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 microstate 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 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 $\Gamma$ whose $3N$ axes 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 $\Gamma$ (irrespectively of whether the system is microscopically isolated or not).

A macroscopic quantity (an observable) which characterizes the 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 measurements made on the system do not really measure instantaneous values of the observable $O$ but rather its time average of the sort

$$O_{\text{obs}} = \frac{1}{\tau} \int_{t_0}^{t_0 + \tau} dt O(q, p), \quad (174)$$

in which $\tau$ is sufficiently long on the microscopic scale to smooth out fluctuations. Obviously, the formula (174) stays valid whether the system is completely isolated or not. In principle, 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

$$\dot{q}^i = \frac{\partial \mathcal{H}(q, p)}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}(q, p)}{\partial q^i}, \quad i = 1, \ldots, 3N, \quad (175)$$

and the initial conditions $q^i(t_0), p_i(t_0)$ set at some instant $t_0$. The measured quantity
(174) could be then in principle computed in terms of these initial data as

\[ O = \frac{1}{\tau} \int_{t_0}^{t_0 + \tau} dt \, O(q(t), p(t)). \]  

(176)

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 \( N \sim N_A = 6.022 \times 10^{20} \) particles is certainly beyond reach. Secondly, and this ultimately turns out to be crucial, 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 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 may be not isolated microscopically or even macroscopically, when it is in contact with e.g., a heat bath) 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 equilibrium as a macroscopically isolated system, or in equilibrium with its environment; slightly 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 \( t_0 \). For example the ensemble representing at the moment \( t_0 \) a real system in equilibrium (either macroscopically isolated or in equilibrium with its environment) whose macrostate does not change in time, could be constructed by taking all points of the phase space \( \Gamma \) which this system “visits” in consecutive instants \( t_0 + n\Delta t \) with some \( \Delta t \) and \( n \to \infty \) and assigning them all as the initial data at the instant \( t_0 \) to members of the ensemble. The systems of the ensemble populate, therefore, different possible microstates which give rise to the same macrostate (as specified by its macroscopically measured characteristics) and can be represented as a collection of points in the phase space \( \Gamma \). However in view of the fact that the construction of the ensemble evidently cannot be carried out in practice, the main task of the statistical approach is to theoretically construct an ensemble in such a way that the distribution of its systems over the phase space \( \Gamma \) (at some \( t_0 \)) is from the macroscopic point of view representative for the considered real system (at the same instant \( t_0 \)).

In general ensembles obtained either in the way specified above or constructed theoretically can be characterized by their phase space distribution function \( \rho(q, p, t) \)
normalized to unity
\[ \int_{\Gamma} d\Gamma_{(q,p)} \rho(q,p,t) = 1, \quad (177) \]
(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 the ensemble average \( \overline{O}(t) \) over the phase space can be associated:

\[ \overline{O}(t) \equiv \int_{\Gamma} d\Gamma_{(q,p)} \rho(q,p,t) O(q,p). \quad (178) \]

It is the ensemble average which in the statistical approach is going to replace the time averages (174) and the justification of this replacement will be discussed (to the extent to which it is possible) below.

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 (175) with the Hamiltonian \( \mathcal{H}(q,p) \) which includes only the internal interactions of the system (the terms \( \mathcal{H}_{\text{env}} \) and \( \mathcal{H}_{\text{syst-env}} \) are absent from it) and defines therefore the dynamics of an autonomous system. As each of the systems of the ensemble evolves in time, the distribution function \( \rho \) globally characterizing their distribution over the phase space changes. Certain global features of these changes of \( \rho \) 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 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 an arbitrarily chosen neighbouring point \( (q^i + dq^i, p_i + dp_i) \) of the phase space at the moment \( t + dt \) written in terms of its value \( \rho \) 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 linked to \((q^i, p_i)\) by the motion of the system. This requires correlating \(dq^i\) and \(dp_i\) with \(dt\) by \(dq^i = \dot{q}^idt, \ dp_i = \dot{p}_idt\). 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 (175) and the definition of the Poisson bracket)

\[
\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left( \frac{\partial \rho}{\partial q^i} \dot{q}^i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) \equiv \frac{\partial \rho}{\partial t} + \{\rho, \mathcal{H}\}_{PB} .
\]  

(179)

The first part of the Liouville theorem states that \(d\rho/dt\) is zero.

To prove this statement we consider an arbitrarily chosen finite (in the mathematical language, of finite measure) domain \(\Delta \Gamma\) of the phase space \(\Gamma\) 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 \((d\mathcal{N}_{\Delta \Gamma}(t)/dt)dt\) 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”

\[
\mathcal{N}_\rho \mathbf{u} \equiv \mathcal{N}_\rho \left( \frac{q^i}{\dot{p}_i} \right) ,
\]

(which is the product of \(\mathcal{N}_\rho\) itself and the \(6N\) dimensional “velocity vector” \(\mathbf{u}\), integrated over the boundary of the domain \(\Delta \Gamma\) and multiplied by \(dt\) \((d\Sigma\) is the \(6N - 1\)-dimensional differential area vector normal to the surface and directed outwards):

\[
d\mathcal{N}_{\Delta \Gamma}(t) = -dt \int_{\partial \Delta \Gamma} d\Sigma \cdot \mathbf{u} \mathcal{N}_\rho .
\]

By the Stokes theorem this can be written in the form

\[
\frac{d\mathcal{N}_{\Delta \Gamma}(t)}{dt} = -\int_{\Delta \Gamma} d\Gamma_{(q,p)} \nabla \cdot (\mathbf{u} \mathcal{N}_\rho) .
\]

Equating the two forms of the rate \(d\mathcal{N}_{\Delta \Gamma}(t)/dt\) and taking into account that the domain \(\Delta \Gamma\) was arbitrary (which implies that the integral equality can be replaced by its local differential version), we obtain the equality

\[
0 = \frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{u} \rho) = \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left[ \frac{\partial \rho}{\partial q^i} (\dot{q}^i \rho) + \frac{\partial \rho}{\partial p_i} (\dot{p}_i \rho) \right]
\]

\[
= \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left[ \frac{\partial \rho}{\partial q^i} \dot{q}^i + \frac{\partial \rho}{\partial p_i} \dot{p}_i + \left( \frac{\partial \dot{q}^i}{\partial q^i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) \rho \right] .
\]

127
Since
\[ \frac{\partial q^i}{\partial q^i} = \frac{\partial}{\partial q^i} \frac{\partial \mathcal{H}}{\partial p_i} = \frac{\partial^2 \mathcal{H}}{\partial q^i \partial p_i}, \quad \frac{\partial p_i}{\partial p_i} = -\frac{\partial}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q^i} = -\frac{\partial^2 \mathcal{H}}{\partial p_i \partial q^i}, \]
the last term in the square brackets is zero and comparing with (179) we indeed conclude that \( d\rho/dt = 0 \). The vanishing of the total time derivative of the distribution function \( \rho \) 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
\[ \rho(q, p, t) = \rho(q_0, p_0, t_0), \quad (180) \]
holding for \((q^i, p_i)\) related at \( t \) to the values \((q^i_0, p^i_0)\) of the phase space coordinates at \( t_0 \) by the motion: \( q^i = q^i(t, q_0, p_0, t_0) \) and \( p_i = p_i(t, q_0, p_0, t_0) \).

To obtain the second part of the Liouville theorem we consider a small (infinitesimal in fact) domain \( \Delta \Gamma \) of the phase space and follow over the time interval \( s \) 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 \( \Delta \Gamma \). Moreover, no one point of those which at \( t_0 \) formed the interior of \( \Delta \Gamma \) could have evolved outside of the new domain, because to do 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 and since the Hamilton’s equations of motion are causal it would have to coincide all the time 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 changed between \( t_0 \) and \( t_0 + s \), its volume remains constant. This in turn means that
\[ J = \frac{\partial(q, p)}{\partial(q_0, p_0)} = 1. \quad (181) \]
where \( J \) is the jacobian of the change (transformation) of the integration variables from \((q^i, p_i)\) to \((q^i_0, p^i_0)\) 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 above argument shows that
\[ \int_{\Delta \Gamma(t_0+s)} d\Gamma_{(q,p)} \mathcal{N} \rho(q, p, t_0 + s) = \int_{\Delta \Gamma(t_0)} d\Gamma_{(q_0,p_0)} \mathcal{N} \rho(q_0, p_0, t_0), \]
where \( q^i = q^i(s, q_0, p_0, t_0) \), \( p_i = p_i(s, q_0, p_0, t_0) \). After using these formulae as defining the transformation of the integration variables and using the equality (180) the left integral takes the form
\[ \int_{\Delta \Gamma(t_0)} d\Gamma_{(q_0,p_0)} \frac{\partial(q, p)}{\partial(q_0, p_0)} \mathcal{N} \rho(q_0, p_0, t_0), \]

128
Its equality to the right integral means therefore that (181) holds.

**Correspondence of 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 (174)

\[
O_{\text{obs}} = \frac{1}{\tau} \int_{t_0}^{t_0 + \tau} dt \, O(q, p),
\]

measured on the real system (isolated or not) and the appropriate ensemble average. The latter is defined as follows. We construct an ensemble which at \( t_0 \) is representative for the real system. With each (isolated) system of the ensemble which from its microstate \((q_i^0, p_i^0)\) at \( t_0 \) evolves as dictated the autonomous Hamilton’s equations (175):

\[
q^{i'} = q^i(s, q_0, p_0), \quad p_i = p_i(s, q_0, p_0),
\]

the time average

\[
O_\tau(q_0, p_0) = \frac{1}{\tau} \int_0^\tau ds \, O(q(s, q_0, p_0), p(s, q_0, p_0)),
\]

can be associated. The period \( \tau \) (here and in (174)) can be taken sufficiently long from the microscopic point of view (for instance, in the case of a gas \( \tau \) should be large compared to the mean time between collisions), so that \( O_\tau \) is practically independent of \( \tau \) (though, as indicated by the notation, it can still depend on the initial point \((q_i^0, p_i^0)\)). The main postulate is then that \( O_{\text{obs}} \) should be identified with the ensemble average

\[
\overline{O_\tau} \equiv \int d\Gamma(q_0, p_0) \rho(q_0, p_0, t_0) O_\tau(q_0, p_0).
\]

The validity of the identification of \( O_{\text{obs}} \) with \( \overline{O_\tau} \) depends of the magnitude of fluctuations of the values of \( O_\tau \)’s around the mean \( \overline{O_\tau} \). If the real system is (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 low, so the same must be true of \( O_{\text{obs}} \). If the real system is not isolated (either microscopically or because it interacts with the 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 the 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 that the probability that \( O_{\text{obs}} \) which is a composition (in varying proportions) of parts of different \( O_\tau \)'s deviates significantly from the ensemble average \( \overline{O_\tau} \) is low if the fluctuations (which are determined by the ensemble) of \( O_\tau \)'s are small provided the ensemble does represent the real syste
faithfully. In either case, the validity of the identification of $O_{\text{obs}}$ with the ensemble average $\overline{O_r}$ is reduced to the problem whether the fluctuations of $O_r$'s around $\overline{O_r}$ 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 the fluctuations are indeed small owing to the smoothing effect of time averaging involved in computing individual $O_r$'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 $\overline{O_r}$ 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 $\overline{O_r}$ of observable quantities $O$ (but not their fluctuations) are to some extent insensitive to the form of $\rho$.

The construction of the statistical ensembles and the prescription for computing the ensemble average $\overline{O_r}$, which according to the adopted postulate is to be identified with the really measured value $O_{\text{obs}}$ given above, are fairly general and apply to (real) systems in equilibrium (either as macroscopically isolated system or with their surrounding) as well 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 $O_{\text{obs}}$ 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 $\overline{O_r}$ given by (182) 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:

$$\frac{\partial}{\partial t} \rho(q, p, t) = 0.$$ (183)

Since Liouville theorem states that $d\rho/dt = 0$, from the formula (179) it follows that the distribution function $\rho$ representing a systems in equilibrium, in addition to being not explicitly dependent on time, must have zero Poisson brackets with the Hamiltonian of the system (treated as isolated):

$$\{\rho(q, p), \mathcal{H}(q, p)\}_{\text{PB}} = 0.$$ (184)

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 a very large numbers of particles have only a few of conserved quantities: 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

---

81This is why the problem of choosing the representative ensemble is the central one of the 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 good.

82The 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.
of an ensemble representing a real system in equilibrium has the form

$$\rho(q, p, t) = \rho(q, p) = \rho(H(q, p)).$$

(185)

Moreover, the recipe for calculating the ensemble average (182) pertaining to the properties of a system in equilibrium can be considerably simplified: in the explicit formula

$$\overline{O}_\tau = \frac{1}{\tau} \int d\Gamma_{(q_0, p_0)} \rho(q_0, p_0) \int_0^\tau ds \, O(q(s, q_0, p_0), p(s, q_0, p_0))$$

$$= \frac{1}{\tau} \int_0^\tau ds \int d\Gamma_{(q_0, p_0)} \rho(q_0, p_0) O(q(s, q_0, p_0), p(s, q_0, p_0)),$$

one can use the result (180) with \( q^i = q^i(s, q_0, p_0), p_i = p_i(s, q_0, p_0) \) specified to distribution functions satisfying (183) and the result (181) 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 \( \tau \)-independent formula

$$\overline{O}_\tau \equiv \overline{O} = \int d\Gamma_{(q, p)} \rho(q, p) O(q, p).$$

(186)

In the case of systems in equilibrium, when taking the limit \( \tau \to \infty \) in the definition (174) of the measured quantity is allowed (because the measured macroscopic properties do not depend on time) the formula (186) can be justified a la Landau: one can in this case imagine that the phase space trajectory of the real system is followed over a period \( \tau \) 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 distribution function is constructed by taking

$$\rho(q, p) \propto \lim_{\tau \to \infty} \frac{\Delta t_{(q, p)}(\tau)}{\tau},$$

(187)

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 right 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 \( \rho \) similar to the one defined here operationally is not solved by the above reasoning.

131
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\(^{83}\)

$$
\rho_O(O) = \int d^\Gamma(q, p) \rho(q, p) \delta(O(q, p) - O) .
\tag{188}
$$

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 (186) or, as

$$
\overline{O} = \int dO O \rho_O(O) .
\tag{189}
$$

The standard estimate of the fluctuations of the random variable $O$ around its mean value
$\overline{O}$ is obtained by computing the quantity

$$
\sigma_O^2 \equiv (O - \overline{O})^2 = \int dO (O - \overline{O})^2 \rho_O(O) = \overline{O^2} - \overline{O}^2 .
\tag{190}
$$

called the mean quadratic fluctuation and the taking its square root. The (dimensionless)
measure of fluctuations is the relative fluctuation given by the ratio

$$
\frac{\sqrt{\sigma_O^2}}{O} ,
\tag{191}
$$

which tells how large fluctuations are compared to the mean value.\(^{84}\) The success of
the statistical method applied to large ($N \sim N_A$) systems relies mostly on the fact that even
if the distribution $\rho(q, p)$ 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 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.

\(^{83}\)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 continuous 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 $O$.

\(^{84}\)Of course, if $\overline{O} = 0$, one has to invent another dimensionless measure of fluctuations.
Statistical independence

If the system (the space of elementary events) can be split into two mutually (almost) noninteracting parts \(a\) and \(b\) 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

\[
d\Gamma_{(q_a,p_a,q_b,p_b)} \rho_{a,b}(q_a, p_a, q_b, p_b, t) = d\Gamma_{(q_a,p_a)} \rho_a(q_a, p_a, t) d\Gamma_{(q_b,p_b)} \rho_b(q_b, p_b, t) .
\] (192)

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.\(^{85}\)

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 molecules. 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 (191) 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,

\[
\sigma^2_O = \overline{(O - \overline{O})^2} = \left( \sum_a (O_a - \overline{O_a}) \right)^2 = \sum_a \sum_b (O_a - \overline{O_a})(O_b - \overline{O_b}).
\]

\(^{85}\)In general the statistical independence holds over not too long periods; in the case of systems in equilibrium whose macroscopic state does 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
\[
\sigma^2_O = \sum_a (O_a - \overline{O}_a)^2 + \sum_{a\neq b} (O_a - \overline{O}_a)(O_b - \overline{O}_b) \\
= \sum_a (O_a - \overline{O}_a)^2 + \sum_{a\neq b} (O_a - \overline{O}_a)(O_b - \overline{O}_b). \tag{193}
\]

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
\[
\frac{\sqrt{\sigma^2_O}}{\overline{O}} \propto \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}, \tag{194}
\]
and is really minute, if the system is large \( (N \sim N_A) \).

Another important consequence of the statistical independence which is implied by the factorization (192) of the distribution function is as follows. Suppose the two subsystems can be treated as statistically independent, either because they are separated one from another or because, as in the considerations above, they are two macroscopic parts of the same macroscopic body. The Hamiltonian of the entire system in equilibrium treated as microscopically isolated (even if the real system is not necessarily such - it can be in equilibrium with its environment) is then just the sum \( \mathcal{H} = \mathcal{H}_a + \mathcal{H}_b \) (\( \mathcal{H}_a \) depends on the variables of the subsystem \( a \), while \( \mathcal{H}_b \) on the variables of the subsystem \( b \)). In view of the statistical independence of the subsystems \( a \) and \( b \), the distribution function of the ensemble representing the entire system must have the property (recall the result (185) !)
\[
\rho_{a,b}(\mathcal{H}) = \rho_{a,b}(\mathcal{H}_a + \mathcal{H}_b) = \rho_a(\mathcal{H}_a) \rho_b(\mathcal{H}_b),
\]
or
\[
\ln \rho_{a,b}(\mathcal{H}_a + \mathcal{H}_b) = \ln \rho_a(\mathcal{H}_a) + \ln \rho_b(\mathcal{H}_b). \tag{195}
\]
The only possibility of satisfying this requirement is
\[
\ln \rho_a(\mathcal{H}_a) = \alpha_a - \beta \mathcal{H}_a, \tag{196}
\]
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),
\[
\ln \rho_a(\mathcal{H}_a) = \alpha_a - \beta \mathcal{H}_a + \gamma N_a. \tag{197}
\]

\[\textit{\textsuperscript{86}}\text{Indeed, differentiating the equality } \ln f(x + y) = \ln f(x) + \ln f(y) \text{ with respect to } x \text{ one obtains that } f'(x + y)/f(x + y) = f'(x)/f(x); \text{ in the same way one finds that } f'(x + y)/f(x + y) = f'(y)/f(y); \text{ hence } f'(y)/f(y) = f'(x)/f(x) = -\beta = \text{ const.}, \text{ because both sides are functions of a different independent variable. From this the conclusion follows.}\]
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. 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 environment, 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

\[
\rho(q, p) = \begin{cases} 
\text{Const.} & \text{if } E \leq \mathcal{H}(q, p) \leq E + \Delta E \\
0 & \text{otherwise}
\end{cases}
\]

(198)

The ensemble defined by this distribution function is called microcanonical. Since the distribution function \(\rho\) must be normalized to unity, the constant is equal \(1/\Gamma(E, V, N, \Delta E)\) where

\[
\Gamma(E, V, N, \Delta E) = \int_{E \leq \mathcal{H}(q, p) \leq E + \Delta E} d\Gamma(q, p),
\]

(199)

is the volume of the corresponding part of the phase space. In various considerations useful are also two other quantities:

\[
\Sigma(E, V, N) = \int_{\mathcal{H}(q, p) \leq E} d\Gamma(q, p) \equiv \int d\Gamma(q, p) \theta(E - \mathcal{H}(q, p)),
\]

(200)

which gives the phase space volume corresponding to the system’s total energy less than \(E\) and:

\[
\omega(E, V, N) = \frac{\partial\Sigma(E, V, N)}{\partial E} = \int d\Gamma(q, p) \delta(\mathcal{H}(q, p) - E),
\]

(201)

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 \(\Delta E\) is small compared to \(E\), a useful working approximation is

\[
\Gamma(E, V, N, \Delta E) \approx \frac{\partial\Sigma(E, V, N)}{\partial E} \Delta E = \omega(E, V, N) \Delta E.
\]

(202)

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\textsuperscript{87}

\[ S_{\text{stat}} = k_B \ln \Gamma(E, V, N, \Delta E), \tag{203} \]

in which \( k_B = 8.617 \times 10^{-5} \text{ eV} \cdot \text{K}^{-1} \) is the Boltzmann constant. 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 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 \( S_{\text{stat}} \) defined by (203) 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 identified with

\[ S = N \lim_{\infty} \left( \frac{S_{\text{stat}}(E, V, N, \Delta E)}{N} \right), \tag{204} \]

where the symbol \( \lim_{\infty} \) denotes the so-called \textbf{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 \( \Delta E \) drops out.\textsuperscript{88} In most cases the statistical entropies computed replacing \( \Gamma(E, V, N, \Delta E) \) under the logarithm in (203) by \( \Sigma(E, V, N) \) or by \( \omega(E, V, N) \) lead in the thermodynamic limit to the same thermodynamic entropy.

Since the statistical entropy (203) 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 no such an ambiguity (apart from the choice of the energy allowance \( \Delta E \) ) in quantum statistical physics in which entropy (in the thermodynamic limit) is assigned unambiguously to each macrostate. It will be seen that if classical statistical mechanics is to be the limit of its quantum counterpart, the measure \( d\Gamma_{(q,p)} \) should have the form

\[ d\Gamma_{(q,p)} = \frac{d^{3N}q d^{3N}p}{N!(2\pi\hbar)^{3N}}. \tag{205} \]

The factor \((2\pi\hbar)^{3N}\) in the denominator makes the measure dimensionless. It is related to the rule (following e.g. from the Bohr-Sommerfeld quantization condition) that in one-dimensional quasi-classical motion of a single particle every single quantum state can be associated with the “area” \( 2\pi\hbar \) of the two-dimensional phase space. Thus

\textsuperscript{87} This is the famous Boltzmann formula which in the form

\[ S = k \ln W, \]

is engraved on his tomb in der Haupstad Wien Friedhof (at the Viener cemetery).

\textsuperscript{88} 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 \( \Delta E \) drops out in the thermodynamic limit in both cases.

136
$d^{3N}q d^{3N}p/(2\pi\hbar)^{3N}$ 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 distinguishable. 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 in indistinguishable, one and the same quantum state. Therefore the number of quantum states available to particles is reduced roughly\(^{89}\) by the factor of $N!$. This explains the origin of the extra factor of $N!$ in the denominator of the measure (205). 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 a one kind and of $N_2$ (indistinguishable) particles of a second kind (the two kinds are different, so distinguishable) the appropriate measure would be

$$d\Gamma_{(q,p)} = \frac{d^{3N_1}q^{(1)} d^{3N_2}p^{(1)} q^{3N_2}q^{(2)} d^{3N}p^{(2)}}{N_1! (2\pi\hbar)^{3N_1} N_2! (2\pi\hbar)^{3N_2}},$$

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 (205) the Boltzmann formula (203) states that entropy (in units of $k_B$) of a macroscopically isolated system is just the logarithm of the number of (microscopic) quantum states consistent with the macroscopic characteristics of (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 (198), the entropy cannot be given by the formula (203) because the volume of the phase space available 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)

---

\(^{89}\)This does not allow to count the number of states correctly 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.
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. We will see that 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 to justify the postulate of a priori equal probabilities directly on the basis of microscopic time evolution of large autonomous isolated dynamical systems.\textsuperscript{90} It is in this context that the ergodic problem appeared.

Briefly, the problem in which the ergodic problem appeared was to show rigorously\textsuperscript{91} that if the macroscopic system is isolated, the time average (174) 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 the point representing the time evolution of an isolated system in the phase space spends in a given domain $\Delta \Gamma$ the time interval $\Delta t$ which is proportional to its volume.\textsuperscript{92} The second part is showing that this is true independently of where in the phase space such a volume is located. If both parts are shown then the equivalence of the (infinite) time averages and the corresponding phase space averages follows just as in the Landau’s reasoning because this means that $\rho(q,p)$ defined as in (187) would be 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 $\Delta \Gamma(t_1)$ as in Figure 13. At a later instant $t_2$ the same points fill the domain $\Delta \Gamma(t_2)$ which is uniquely determined by the positions at $t_2$ of those phase points which at $t_1$ formed the boundary of $\Delta \Gamma(t_1)$. By the Liouville theorem the volumes of the two domains are equal (though their shapes may be different). Let us follow the points $A$ and $B$ which lie on the same trajectory. 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 $\Delta \Gamma(t_1)$ and $\Delta \Gamma(t_2)$. Moreover

$$t_{AC} = t_{AB} + t_{BC}, \quad t_{BD} = t_{BC} + t_{CD},$$

from which it follows that $t_{AB} = t_{CD}$: the phase point does indeed spend the same times

\textsuperscript{90}At least those who were not contented with the results the postulate leads to as its ultimate justification.

\textsuperscript{91}And this is what attracted pure mathematicians to this field so that it effectively became a branch of mathematics.

\textsuperscript{92}We will simplify slightly the problem and will consider the $6N$-dimensional domains assuming implicitly that the limit $\Delta E \to 0$ is taken; otherwise we would have to consider the projections of the measure $d\Gamma(q,p)$ onto the hypersurface of constant energy.
in $\Delta \Gamma(t_1)$ and $\Delta \Gamma(t_2)$ of equal volumes. Furthermore, if $\tau$ is sent to infinity, any time the phase point passes through $\Delta \Gamma(t_1)$ it must also pass through $\Delta \Gamma(t_2)$.

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 interesting as a problem in pure mathematics and also as 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 removing 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. In the mathematical formulation quantum states are represented in a Hilbert space $\mathcal{H}$ - a complete (in the sense of convergence 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 $\Psi$ and $\Psi' = e^{i\theta}\Psi$ 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

$$ (\Phi|\hat{O}^\dagger\Psi) = (\hat{O}\Phi|\Psi), \quad (206) $$

which must hold for all vectors $\Psi$ and $\Phi$ belonging to $\mathcal{H}$. The spectrum - the set of its eigenvalues $\omega_k$ i.e. the set of such numbers that there exists vectors satisfying the equation $\hat{O}\Psi_k = \omega_k\Psi_k$ - 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 that operator. The most important operator is the (self-adjoint) Hamiltonian $\hat{H}$ representing energy of the system, which determines the system’s state time evolution via the Schrödinger equation

$$ i\hbar \frac{d}{dt} \Psi(t) = \hat{H}\Psi(t). \quad (207) $$

If the Hamiltonian operator is not explicitly time-dependent, the solution of the Schrödinger equation is given by

$$ \Psi(t) = \exp \left( -\frac{i}{\hbar} \hat{H} (t - t_0) \right) \Psi(t_0), \quad (208) $$

\(^{93}\)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 environment - evidently could not liberate himself from the conventional notion of the “wave function” understood narrowly in the spirit of the Schrödinger wave mechanics.
with the state-vector $Ψ(t_0)$ playing the role of the initial conditions. The spectrum of the Hamiltonian which in general can consist of a discrete part and (if the quantum mechanics is formulated in the infinite space) of a continuous part - the corresponding eigenvectors are nonnormalizable generalized vectors belonging to $ℋ^*$ - constitutes always the most important characteristics of the quantum system.

Wherever more subtle mathematical issues do not intervene, convenient is the Dirac notation in which state-vectors are written as “kets” $|Ψ⟩$ and elements of the dual space as $⟨Φ|$. The scalar product $⟨Φ|Ψ⟩$ takes in this notation the form (which in fact implicitly employs the Frechet-Riesz isomorphism) $⟨Φ|Ψ⟩$ and the action of an operator $O$ on a state vector $Ψ$ is written as $O|Ψ⟩$. The quantities $⟨Φ|O|Ψ⟩$ are called matrix elements\(^4\) of the operator $O$ between the states $Ψ$ and $Φ$. As is usually the case with vector spaces, also in the Hilbert space it is possible to chose a basis, a set of vectors $Ψ_l$, or just $|l⟩$, labeled by some label $l$; the basis of the proper Hilbert space $ℋ$ can always be chosen so that the label $l$, which can also stand for a multi-label of the sort $l_1l_2...l_n$, is discrete. Only calculational convenience dictates employing continuos labels (the corresponding vectors are then generalized nonnormalizable vectors, i.e. elements of $ℋ^*$). If a countable basis can be chosen (so that $l$ runs over a finite or a countably infinite set of values), the Hilbert space $ℋ$ is separable. If it is uncountable - e.g. when $l = l_1l_2...l_n$ with $n = ∞$ and each $l_i$ runs over a countably infinite set - then $ℋ$ 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 at the basis of several phenomena (Bose-Einstein condensation, phase transitions), but we will not enter into these fine mathematical details in this Course. Any state-vector can be written as a linear combination

$$Φ = \sum_l Ψ_l c_l, \quad \text{or} \quad |Φ⟩ = \sum_l |l⟩ c_l,$$

of vectors $Ψ_l$ or $|l⟩$ forming the chosen basis. Usually one works with bases formed by mutually orthonormal vectors: $⟨Ψ_ℓ|Ψ_l⟩ = ⟨l'|l⟩ = δ_{l'l}$, which are eigenvectors of an observable $O$ (in most cases of the Hamiltonian $ᴴ$ of the system under consideration). Matrix elements of any operator $O$ between the state-vectors $|l⟩$ of the basis formed by its own eigenvectors are “diagonal” that is, take the form

$$⟨l'|O|l⟩ = o_l δ_{l'l},$$

and the operator itself can be written in the form

$$O = \sum_l |l⟩ o_l ⟨l|= \sum_{o_l ∈(spectrum)} o_l ˆP_{o_l},$$

where $ˆP_{o_l} = ˆP_{o_l}^2$ is the projector onto the subspace (which may well be multidimensional) of $ℋ$ corresponding to the eigenvalue $o_l$ of the operator $O$. Also important is the fact that

\(^4\)The Dirac notation is adapted to self adjoint operators; if $O$ is not self adjoint one has to adopt the convention that the symbol $⟨Φ|O|Ψ⟩$ means $⟨Φ|OΨ⟩$. The scalar product $⟨Φ|O|Ψ⟩$ can in this notation be written only as $((Ψ|O|Φ))^∗$. 

141
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 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 environments). 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 formulation 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 quantum systems which are macroscopic (i.e. consist of $N \sim N_A$ elements - particles or molecules). 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.\(^{95}\)

Yet the spectra of Hamiltonians of macroscopic systems are so dense that they are nearly continuous. Roughly, the energy levels of such systems are separated by gaps of order

$$E_{n+1} - E_n \sim \exp(-N).$$

This is not seen if the $N$ elements of which the system is composed are mutually noninteracting (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 cause smearing

\(^{95}\)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, because in the finite volume the entire spectrum is discrete and the density of states near $E = 0$ is also 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 calculational 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 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 $\mathbb{R}^1$, however large, and all quantities like transmission and reflection coefficients must be replaced by other characteristics simply because the asymptotic (scattering) states cannot be in this case defined.
of degenerate energy levels so that the above estimate of typical interlevel gaps 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 is interpretable in terms of individual energy levels of separate particles. In view of this extreme narrowness of gaps between energy levels of a macroscopic system, any external perturbation, however weak, is associated with the energy transfer to or from the system which is much much larger than its energy gaps and therefore the system, even if it macroscopically seems isolated, is continuously making innumerable transitions between 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 $\Delta E$ takes at least time $\Delta t$ related to $\Delta E$ by

$$\Delta E \Delta t \geq \hbar.$$  

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. $\Delta E$ smaller than typical energy splittings) would require time longer than the lifetime of the Universe.

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 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 that the world is divided into the considered system and its environment. Even though elements which constitute the system and those which constitute the environment can be separately identified and therefore the Hilbert space $\mathcal{H}$ of the entire universe (consisting of the system and the environment) can be represented as the tensor product $\mathcal{H} = \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{env}}$ which in particular means that as the basis of $\mathcal{H}$ one can take vectors

$$|l, L\rangle \equiv |l\rangle \otimes |L\rangle,$$

where $|l\rangle \in \mathcal{H}_{\text{sys}}$ and $|L\rangle \in \mathcal{H}_{\text{env}}$. Quantum states of both, of the system and the environment, i.e. of the Universe, are usually highly entangled, that is, state-vectors representing in $\mathcal{H}$ the state of the universe must be written as general superpositions

$$|\Psi\rangle = \sum_l \sum_L |l\rangle \otimes |L\rangle c_{L,l},$$

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 $c_{L,l} = c_{L,l}^{\text{env}} c_{L,l}^{\text{sys}}$).

Suppose now that in the Hamiltonian of the Universe $\hat{H} = \hat{H}_{\text{sys}} + \hat{H}_{\text{env}} + \hat{H}_{\text{int}}$ in which $\hat{H}_{\text{sys}} \equiv \hat{H}_{\text{sys}} \otimes 1_{\text{env}}$ and $\hat{H}_{\text{env}} \equiv 1_{\text{sys}} \otimes \hat{H}_{\text{env}}$ (which means that they act essentially each in only one space of the tensor product) the interaction term is small and we are interested in an observable measured on the system alone and not on the environment.
Such an observable is represented, therefore, by an operator of the sort \( \hat{O} \equiv \hat{O}_{\text{sys}} \otimes \hat{\imath}_{\text{env}} \).

We would like to express a quantum mechanical expectation value of \( \hat{O} \) in a state of the form (213) entirely by objects pertaining to \( \mathcal{H}_{\text{sys}} \). This can be done with the help of the **density operator** which is introduced in the following way. We start by writing

\[
\langle \Psi | \hat{O} | \Psi \rangle = \left( \sum_{l',l} c^*_{L,l'} \langle L' | \otimes | l' \rangle \right) \hat{O}_{\text{sys}} \left( \sum_{l} | l \rangle \otimes | L \rangle \right) c_{L,l} = \sum_{l',l} \langle l'| \hat{O}_{\text{sys}} | l \rangle \sum_{L} c_{L,l} c^*_{L,l'} . \tag{214}
\]

We have used the way \( \hat{O} \) acts on the state-vectors of the product basis of the complete Hilbert space \( \mathcal{H} \):

\[
\hat{O}(| l \rangle \otimes | L \rangle) = \hat{O}_{\text{sys}} \otimes \hat{\imath}_{\text{env}} (| l \rangle \otimes | L \rangle) = (\hat{O}_{\text{sys}} | l \rangle) \otimes (\hat{\imath}_{\text{env}} | L \rangle) = (\hat{O}_{\text{sys}} | l \rangle) \otimes | L \rangle ,
\]

and the orthogonality\(^{96}\) \( \langle L'|L \rangle = \delta_{L'L} \) of the basis of \( \mathcal{H}_{\text{env}} \). The statistical operator \( \hat{\rho} \) acting in \( \mathcal{H}_{\text{sys}} \) is now defined by giving its matrix elements between the basis vectors \( | l \rangle \) of \( \mathcal{H}_{\text{sys}} \):

\[
(\Psi_l | \hat{\rho} | \Psi_{l'}) = \langle l | \hat{\rho} | l' \rangle = \sum_{L} c_{L,l} c^*_{L,l'} . \tag{215}
\]

Using this operator the expectation value (214) can be written in the form

\[
\langle \Psi | \hat{O} | \Psi \rangle = \sum_{l'l'} \langle l | \hat{\rho} | l' \rangle \langle l'| \hat{O}_{\text{sys}} | l \rangle = \text{Tr}(\hat{\rho} \hat{O}_{\text{sys}}) , \tag{216}
\]

in which the trace on the right hand side is restricted to \( \mathcal{H}_{\text{sys}} \). Defined in this way the operator \( \hat{\rho} \) acting in \( \mathcal{H}_{\text{sys}} \) is Hermitian (self-adjoint). To see this, it suffices to apply the definition (206) to the matrix elements of \( \hat{\rho}^\dagger \) between the basis vectors \( | l \rangle \), or \( \Psi_l \):

\[
(\Psi_{l'} | \hat{\rho}^\dagger | \Psi_l) = (\hat{\rho} \Psi_{l'} | \Psi_l) = (\Psi_l | \hat{\rho} \Psi_{l'})^* = \left( \sum_{L} c_{L,l} c^*_{L,l'} \right)^* = \sum_{L} c^*_{L,l} c_{L,l'} ,
\]

and to compare this with the matrix elements of \( \hat{\rho} \) given by (215):

\[
(\Psi_{l'} | \hat{\rho} | \Psi_l) = \langle l'| \hat{\rho} | l \rangle = \sum_{L} c^*_{L,l} c_{L,l'} .
\]

The two right hand sides are identical, which shows that \( \hat{\rho} = \hat{\rho}^\dagger \).

\(^{96}\)The scalar product of 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

\[
\langle \phi'_{\text{env}} | \otimes \langle \psi'_\text{sys} | = \langle \phi'_{\text{env}} | \otimes \langle \psi'_\text{sys} | = \langle \phi'_{\text{env}} | \otimes \langle \psi'_\text{sys} | = \langle \phi'_{\text{env}} | \otimes \langle \psi'_\text{sys} | \cdot | \psi'_{\text{sys}} | \psi_{\text{sys}} \rangle ,
\]

and extended using linearity to all vectors of \( \mathcal{H} = \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{env}} \).
As every self-adjoint operator, \( \hat{\rho} \) can be diagonalized, that is there exists in \( \mathcal{H}_{sys} \) 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

\[
(\Phi_k | \hat{\rho} | \Phi_{k'}) = \langle k | \hat{\rho} | k' \rangle = w_k \delta_{k,k'}
\]

(217)

If one now takes the expectation value (214) of the unit operator \( \hat{1} = \hat{1}_{sys} \otimes \hat{1}_{env} \) so that obviously \( \langle \Psi | \hat{1} | \Psi \rangle = 1 \) and writes the right hand side of the formula (214) using the vectors \( |k\rangle \) as the basis of \( \mathcal{H}_{sys} \) (instead of \( |l\rangle \)), one will get the equality

\[
1 = \langle \Psi | \hat{1} | \Psi \rangle = \text{Tr}(\hat{\rho}) = \sum_k \langle k | \hat{\rho} | k \rangle = \sum_k w_k .
\]

(218)

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\(^{97}\)

\[
\hat{\rho} = \sum_k |k\rangle w_k \langle k| ,
\]

(219)

and using this representation to write in two ways \( \text{Tr}(\rho \hat{O}_{sys}) \) with the particular operator \( \hat{O}_{sys} = |k\rangle \langle k| \), i.e. a self-adjoint projector onto the \( \hat{\rho} \) eigensubspace corresponding to the eigenvalue \( w_k \). On one hand then

\[
\text{Tr}(\hat{\rho} \hat{O}_{sys}) = \sum_{k''} \langle k'' | \left( \sum_{k'} |k'\rangle w_{k'} \langle k'\rangle \right) (|k\rangle \langle k|) | k'' \rangle = w_k ,
\]

and on the other, using the definition (215),

\[
\text{Tr}(\hat{\rho} \hat{O}_{sys}) = \sum_{l',l} \langle l | \hat{\rho} | l' \rangle \langle l'| \hat{O}_{sys} | l \rangle = \sum_L \sum_{l',l} c_{L,l}^* c_{L,l'} \langle l' | k \rangle \langle k | l \rangle = \sum_L \left| \sum_l c_{L,l} \langle k | l \rangle \right|^2 \geq 0 .
\]

Thus, \( 0 \leq w_k \leq 1 \). Finally, written in the basis of the density operator eigenvectors \( |k\rangle \) the formula (216) takes the form

\[
\langle \Psi | \hat{O} | \Psi \rangle = \text{Tr}(\hat{\rho} \hat{O}_{sys}) = \sum_k w_k \langle k | \hat{O}_{sys} | k \rangle .
\]

(220)

It follows that if the system interacting with the environment is considered separately, it cannot be said to be 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

\[^{97}\text{Indeed,}\]

\[
\langle k'' | \hat{\rho} | k' \rangle = \langle k'' | \left( \sum_k |k\rangle w_k \langle k| \right) | k' \rangle = \sum_k \langle k'' | k \rangle w_k \langle k | k' \rangle = w_k \delta_{k,k''} .
\]

145
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 $\mathcal{H}_{sys}$ which are
eigenvectors of $\hat{\rho}$. The expectation value of an operator $\hat{O}_{sys}$ in the pure state $|\psi\rangle$ reads

$$
\langle \psi | \hat{O}_{sys} | \psi \rangle = \sum_{k} c_k^* c_k \langle k | \hat{O}_{sys} | k \rangle = \sum_{k} |c_k|^2 \langle k | \hat{O}_{sys} | k \rangle + \sum_{k \neq k'} c_k^* c_{k'} \langle k' | \hat{O}_{sys} | k \rangle .
$$

(221)

The first term resembles the formula (220) if the factors $|c_k|^2$ are identified\(^{98}\) with $w_k$. But the expectation value of $\hat{O}_{sys}$ in the pure state $|\psi\rangle$ involves also the second term
which has no counterpart in (220). One may say that while a pure state $|\psi\rangle$ is always a
coherent superposition of basis states $|k\rangle$ with definite relative phases 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 (221) goes over
into (220) upon averaging (with flat probability distribution in the $[0, 2\pi]$ range) over
the phases $\varphi_k$ 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). 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

$$
\hat{\rho}^2 = \hat{\rho},
$$

(222)

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 which interacts with its environment, 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 (219) at some initial
moment $t_0$. Of course, as the state of the entire universe evolves with time, the density
operator $\hat{\rho}$ representing in $\mathcal{H}_{sys}$ the real system changes too but, 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 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 interacting with the
rest of the universe is a collection of $\mathcal{N}$ absolutely isolated systems each in one of the
pure states $|k\rangle$. To be representative for the real system, the relative numbers $\mathcal{N}_k$

\(^{98}\)As $w_k \leq |c_k|^2 \leq 1$. 146
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 $t_0 = 0$) are such that (the limits $N_k \to \infty$ are understood)

$$N_k/N = w_k, \quad \text{where} \quad N = \sum_k N_k.$$ 

The mean values over the ensemble of expectation values of operators representing observables are then formally identical to the formula (220):

$$\overline{O} = \frac{1}{N} \sum_k N_k \langle k|\hat{O}|k\rangle = \sum_k w_k \langle k|\hat{O}|k\rangle,$$

and can be written as

$$\overline{O} = \text{Tr}(\hat{\rho}\hat{O}_{\text{sys}}),$$

with the statistical operator $\hat{\rho}$ which at $t_0$ is formally identical with the density operator $\hat{\rho}$ of the real system.\(^9\) 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 in thermal equilibrium with its surrounding, etc.).

Because the systems forming the ensemble are isolated, their time evolution is completely determined by the Hamiltonian $\hat{\mathcal{H}}$ (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 (208):

$$|k\rangle \to |k(t)\rangle = \exp(-(i/\hbar)\hat{\mathcal{H}}t)|k\rangle.$$

as a result

$$\hat{\rho} \to \hat{\rho}(t) = \sum_k w_k e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}|k\rangle\langle k + \frac{i}{\hbar}\hat{\mathcal{H}}t.}$$

Differentiating both sides with respect to $t$ gives the differential equation

$$\frac{d\hat{\rho}(t)}{dt} = -\frac{i}{\hbar} [\hat{\mathcal{H}}, \hat{\rho}(t)],$$

which in quantum statistical mechanics plays the same role as does the Liouville equation

$$\frac{\partial \hat{\rho}(q,p,t)}{\partial t} = \{\mathcal{H}(q,p), \rho(q,p,t)\}_\text{PB},$$

\(^9\)Because of this formal identity one can take the position that the ensemble is just a single system in an appropriate mixed state.
in classical statistical physics.\footnote{The equation (225) looks similar to the Heisenberg equation
\[ \frac{d\hat{O}_H(t)}{dt} = \frac{i}{\hbar} [\hat{H}, \hat{O}_H(t)], \]
satisfied by time-dependent Heisenberg picture operators $\hat{O}_H(t)$ which are counterparts of time-independent Schrödinger picture operators (for more details, see Chapter 1 of my notes to quantum field theory), but the sign of the right hand sides are different; in fact the Heisenberg picture counterpart $\hat{p}_H$ of the statistical operator $\hat{p}(t)$ is time independent. It is also worth noting that the transition from the classical Liouville equation (226) to the equation (225) 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
\[ \{\cdot, \cdot\}_{PB} \rightarrow -\frac{i}{\hbar} [\cdot, \cdot]. \]}

If the real system is in equilibrium (either as a macroscopically isolated system or as a system in equilibrium with its surroundig), which means that macroscopic observables measured on it do not depend on time, the statistical operator of the representative ensemble should be time independent, just as were the distribution functions $\rho(q,p)$ corresponding to classical systems in equilibrium (only then the ensemble mean values (224) will be stationary). From the equation (225) it follows that $d\hat{\rho}/dt = 0$ implies
\[ [\hat{H}, \hat{\rho}] = 0. \tag{227} \]

According to the usual rules of quantum mechanics this means that there exist in $\mathcal{H}$ a basis $|n\rangle$ in which both operators, $\hat{H}$ and $\hat{\rho}$, are simultaneously diagonal:
\[ \hat{H} = \sum_n E_n |n\rangle \langle n|, \quad \hat{\rho} = \sum_n w_n |n\rangle \langle n|. \tag{228} \]

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 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 Hamiltonian is $\hat{H} = \hat{H}_1 \otimes 1_2 + 1_1 \otimes \hat{H}_2$ (the term $\hat{H}_\text{int}$ 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 the product of observables pertaining to different systems, equals $\overline{O_1 O_2}$:
\[ \overline{O} = \overline{O_1 O_2} = \text{Tr}_\mathcal{H} \left( \hat{\rho}_1 \otimes \hat{\rho}_2 \hat{O}_1 \otimes \hat{O}_2 \right). \]

\footnote{The equation (225) looks similar to the Heisenberg equation
\[ \frac{d\hat{O}_H(t)}{dt} = \frac{i}{\hbar} [\hat{H}, \hat{O}_H(t)], \]
satisfied by time-dependent Heisenberg picture operators $\hat{O}_H(t)$ which are counterparts of time-independent Schrödinger picture operators (for more details, see Chapter 1 of my notes to quantum field theory), but the sign of the right hand sides are different; in fact the Heisenberg picture counterpart $\hat{p}_H$ of the statistical operator $\hat{p}(t)$ is time independent. It is also worth noting that the transition from the classical Liouville equation (226) to the equation (225) 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
\[ \{\cdot, \cdot\}_{PB} \rightarrow -\frac{i}{\hbar} [\cdot, \cdot]. \]}

\[ 148 \]
\[
\begin{align*}
&= \text{Tr}_\mathcal{H} \left( \hat{\rho}_1 \hat{O}_1 \otimes \hat{\rho}_2 \hat{O}_2 \right) = \sum_{n_1 n_2} \langle n_2, n_1 | \hat{\rho}_1 \hat{O}_1 \otimes \hat{\rho}_2 \hat{O}_2 | n_1, n_2 \rangle \\
&= \sum_{n_1} \langle n_1 | \hat{\rho}_1 \hat{O}_1 | n_1 \rangle \sum_{n_2} \langle n_2 | \hat{\rho}_2 \hat{O}_2 | n_2 \rangle = \text{Tr}_{\mathcal{H}_1} \left( \hat{\rho}_1 \hat{O}_1 \right) \text{Tr}_{\mathcal{H}_2} \left( \hat{\rho}_2 \hat{O}_2 \right) = \overline{O}_1 \overline{O}_2.
\end{align*}
\]

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:

\[
w_n^{(12)} \delta_{n'n} = \langle n'| \hat{\rho} | n \rangle = \langle n_1', n_2' | \hat{\rho}_1 \otimes \hat{\rho}_2 | n_1, n_2 \rangle = \langle n_1'| \hat{\rho}_1 | n_1 \rangle \langle n_2'| \hat{\rho}_2 | n_2 \rangle = w_n^{(1)} \delta_{n_1'n_1} w_n^{(2)} \delta_{n_2'n_2}.
\]

In other words, \( w_n^{(12)} = w_n^{(1)} w_n^{(2)} \), or

\[
\ln w_n^{(12)} \equiv \ln w_n^{(1)} + \ln w_n^{(2)},
\]

which, since energies of the two (noninteracting or interacting negligibly weakly with one another) parts of the system are additive, \( E_{n_1 n_2} = E_{n_1} + E_{n_2} \) and in general the probabilities can be viewed as functions of energies: \( w_{n_i} = w(E_{n_i}) \), means that

\[
\ln w(E_{n_i}) = \alpha - \beta E_{n_i}, \quad \text{and} \quad \ln w(E_{n_a}) = \alpha_a - \beta E_{n_a},
\]

as this is the only way of satisfying the requirement (229) if \( E_n \equiv E_{n_1 n_2 ...} = E_{n_1} + E_{n_2} + ... \). All this can be concisely written in the form\(^{101}\)

\[
\ln \hat{\rho}^{(12)} = \left( \ln \hat{\rho}^{(1)} \right) \otimes \hat{1}_2 + \hat{1}_1 \otimes (\ln \hat{\rho}^{(2)}).
\]

This imposes a strong constraint on statistical operators of macroscopic systems and will be crucial in deriving the general formula for entropy.

The quantum microcanonical ensemble

We now consider a real system which macroscopically can be treated as isolated from its surrounding and in equilibrium. 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

\[
\hat{\rho} = \text{Const.} \sum_n |n\rangle \langle n|,
\]

\[
E \leq E_n \leq E + \Delta E
\]

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 (232) it has been taken into account that as

\(^{101}\)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.
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 $\Delta E$ (of course $\Delta E \ll E$, in the case of a large system). Of course, since $\text{Tr}(\hat{\rho}) = 1$, the constant in the definition (232) must be given by (the ellipses stand for other macroscopic variables like magnetization, etc. which characterize the system)

$$\text{Const.}^{-1} = \left( \frac{\text{number of quantum states in the interval } [E, E + \Delta E]}{\text{volume}} \right) \equiv \Gamma(E, V, \ldots, N, \Delta E). \quad (233)$$

Since 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. Together the formulae (232) and (233) define the quantum microcanonical ensemble which should be representative for systems macroscopically isolated and in equilibrium. The form (232) of the statistical operator $\hat{\rho}$ is, of course consistent with the general requirements (227) and (230).

As in the case of the classical microcanonical ensemble, the statistical entropy of a macroscopically isolated system in equilibrium is defined to be

$$S_{\text{stat}} = k_B \ln \Gamma(E, V, \ldots, N, \Delta E). \quad (234)$$

As there, it is a measure of disorder of a given equilibrium macrostate quantified in terms of the (logarithm of the) number of quantum microstates realizing that macrostate (i.e. in the Callenian language, consistent with macroscopic constraints the system is subjected to). Notice, however, that in contrast to the classical case, the quantity $\Gamma(E, V, \ldots, N, \Delta E)$ is within 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. As has been explained, the arbitrariness in the classical case is removed by requiring that results of the classical and quantum approaches 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, to define the density $\omega(E, V, \ldots, N)$ of states (around $E_n = E$) by

$$\omega(E, V, \ldots, N) = \frac{\partial \Sigma(E, V, \ldots, N)}{\partial E}, \quad (235)$$

so that the relation (202) between $\Gamma(E, V, \ldots, N, \Delta E)$, $\omega(E, V, \ldots, N)$ and $\Delta E$ holds true also in the quantum case. Then, again as in the classical case, one can define two other

\textsuperscript{102}Here it is important that the theory is formulated in the finite volume: for instance, to the range $[-\varepsilon, 0]$ (with arbitrarily small positive $\varepsilon$) there correspond infinitely many discrete energy eigenstates of the Hydrogen atom Hamiltonian (recall that $E_{nlm, s_z} = -\hbar c^2 \alpha_{EM}^2 / 2n^2$ 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).
statistical entropies replacing in (234) \( \Gamma(E, V, \ldots, N, \Delta E) \) either by \( \Sigma(E, V, \ldots, N) \) or by the density \( \omega(E, V, \ldots, N) \) of states. In the thermodynamical limit, if the energy spectrum of the system is typical, all the three definitions of \( S_{\text{stat}} \) lead to the same thermodynamical entropy

\[
S_{\text{TMD}} = N \lim_{\infty} \left( \frac{S_{\text{stat}}}{N} \right),
\]

or, in cases in which the number of particles cannot be defined (e.g. because the system is a relativistic quantum field)

\[
S_{\text{TMD}} = V \lim_{\infty} \left( \frac{S_{\text{stat}}}{V} \right).
\]

Thermodynamical entropy obtained in this way in the framework of the Microcanonical Ensemble is given as a function of its natural variables 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 (234) have (in the thermodynamical limit) the properties of the entropy defined within 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 (203) and (234) 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 (203) and (234) 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 \( E \) defined on the space of elementary events. In the classical case it is defined using the prescription (188) which here, because when the considered system is in equilibrium \( \rho(q,p) = \rho(\mathcal{H}(q,p)) \), gives

\[
\rho_E(E) = \int d\Gamma(q,p) \rho(\mathcal{H}(q,p)) \delta(\mathcal{H}(q,p) - E)
\]

151
\[ \rho(E) = \rho(E) \int d\Gamma(q,p) \delta(\mathcal{H}(q,p) - E) = \rho(E) \omega(E). \] (238)

In the quantum case one can define the distribution \( \rho_E(E) \) by the equality

\[ \rho_E(E) \, dE = \sum_n w_n, \] (239)

\[ E \le E_n \le E + dE \]

Since \( \sum_n w_n = 1 \), the distribution \( \rho_E(E) \) defined in this way is normalized: \( \int dE \rho_E(E) = 1 \). On the other hand, 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

\[ \rho_E(E) = w(E) \omega(E). \] (240)

The two expressions: the classical one (238) and the quantum one (240) become therefore identical if \( w(E) \) is identified with \( \rho(E) \).

If the considered system is large, the distribution \( \rho_E(E) \) is nonzero only in the close vicinity of a value \( E^* \) which is practically the same as the mean \( \bar{E} \) computed using \( \rho_E(E) \). In other words, \( \rho(E) \) has a Dirac delta-like peak practically at \( E = \bar{E} \). It is therefore possible to characterize the microscopic spread of an 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” \( \Delta E \) of its energy distribution defined by the relation

\[ \rho_E(\bar{E}) \Delta E = 1. \] (241)

This allows to define the effective number of microstates (that is, the number of those microstates which are really relevant in the ensemble) \( \Gamma(\bar{E}, V, N, \Delta E) \) by the equality (in the classical case replace \( w(E) \) by \( \rho(E) \))

\[ 1 = w(\bar{E}) \omega(\bar{E}) \Delta E = w(\bar{E}) \Gamma(\bar{E}, V, N, \Delta E), \] (242)

and then the statistical entropy by

\[ S_{stat} = k_B \ln \Gamma(\bar{E}, V, N, \Delta E) = k_B \ln [\omega(\bar{E}) \Delta E]. \] (243)

It is easy to see that in the case of the mirocanonical statistical operator (232) or the distribution function (198) the definitions (244) and (203) coincide with (243). Indeed, to consider only the quantum case, in this case \( w(E) = w(\bar{E}) = 1/\Gamma(E, V, N, \Delta E) \), which implies the equality of \( \Gamma(\bar{E}, V, N, \Delta E) \) and \( \Gamma(E, V, N, \Delta E) \).

Exploiting now the definition (242) of the effective number of microstates, the formula (243) for entropy can be cast in the form

\[ S_{stat} = -k_B \ln w(\bar{E}), \] (244)

\[ \text{152} \]
It has been argued, however, that statistical independence of subsystems implies that the logarithm of \( w(E) \) must be at most a \textbf{linear} function of energy, \( \ln w(E) = \alpha - \beta E \) (\( \ln \rho(\mathcal{H}) = \alpha - \beta \mathcal{H} \), classically). This means that
\[
\ln w(E) = \overline{\ln w(E)},
\]
which allows to write the formula (243) in the form
\[
S_{\text{stat}} = -k_B \text{Tr}(\hat{\rho} \ln \hat{\rho}), \quad (245)
\]
in the quantum case and as
\[
S_{\text{stat}} = -k_B \int d\Gamma_{(q,p)} \rho(q,p) \ln \rho(q,p) \quad (246)
\]
in the classical case. It is clear that in the case of the classical microcanonical ensemble, when \( \rho(q,p) = \text{Const.} = 1/\Gamma(E, \Delta E) \), and
\[
\int d\Gamma_{(q,p)} \rho(q,p) \ln \rho(q,p) = \frac{1}{\Gamma(E, \Delta E)} \ln \frac{1}{\Gamma(E, \Delta E)} \int_{E \leq \mathcal{H} \leq E + \Delta E} d\Gamma_{(q,p)};
\]
(246) reduces to (203). An analogous reasoning readily shows that (245) reduces to (234).

Expressions (245) and (246) 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 environment, is the equipartition of its mean energy between its microscopic degrees of freedom. Here we demonstrate it using the classical microcanonical ensemble distribution function (198).

We begin by computing the ensemble average

$$q^i \frac{\partial \mathcal{H}}{\partial q^j} = \frac{1}{\Gamma(E, \Delta E)} \int_{E \leq \mathcal{H} \leq E + \Delta E} d\Gamma_{(q,p)} q^i \frac{\partial \mathcal{H}}{\partial q^j} \approx \frac{\Delta E}{\Gamma(E, \Delta E)} \frac{\partial}{\partial E} \int_{\mathcal{H} \leq E} d\Gamma_{(q,p)} q^i \frac{\partial \mathcal{H}}{\partial q^j} .$$

We have used the standard trick

$$\int_{E \leq \mathcal{H} \leq E + \Delta E} d\Gamma_{(q,p)} f(q,p) = \int_{\mathcal{H} \leq E + \Delta E} d\Gamma_{(q,p)} f(q,p) - \int_{\mathcal{H} \leq E} d\Gamma_{(q,p)} f(q,p) = F(E + \Delta E) - F(E) \approx \Delta E \frac{\partial}{\partial E} F(E) .$$

This can be also written as

$$q^i \frac{\partial \mathcal{H}}{\partial q^j} \approx \frac{\Delta E}{\Gamma(E, \Delta E)} \frac{\partial}{\partial E} \int_{\mathcal{H} \leq E} d\Gamma_{(q,p)} \left\{ \frac{\partial}{\partial q^i} \left[ q^i (\mathcal{H} - E) \right] - \delta^i_j (\mathcal{H} - E) \right\} .$$

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 $\mathcal{H} = E$, vanishes the expression under the derivative.\(^{103}\) 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

$$q^i \frac{\partial \mathcal{H}}{\partial q^j} \approx \frac{\Delta E}{\Gamma(E, \Delta E)} \int_{\mathcal{H} \leq E} d\Gamma_{(q,p)} \delta^i_j = \frac{\Delta E}{\omega(E) \Delta E} \delta^i_j \Sigma(E) = \frac{\Sigma(E)}{\omega(E)} \delta^i_j .$$

\(^{103}\)More precisely, the domain of the integrations over the variables $q^i$ should be effectively restricted by introducing a smooth boundary potential $\nu(q^1, \ldots, q^{3N})$ 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

$$\nu \to \begin{cases} 0 & \text{if all } q^i \in V, \\ \infty & \text{otherwise} \end{cases}$$

should be taken.
(The relation (202) has been used here.) Using the relation \( \omega(E) = \partial \Sigma(E)/\partial E \), this result can also be written as

\[
q^i \frac{\partial \mathcal{H}}{\partial q^j} = \delta^i_j \left( \frac{\partial \ln \Sigma(E)}{\partial E} \right)^{-1}.
\]

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

\[
q^i \frac{\partial \mathcal{H}}{\partial q^j} = -q^i \dot{p}_j = k_B T \delta^i_j.
\]  \hspace{1cm} (247)

Exactly analogous computation\(^{104}\) gives the result

\[
p_i \frac{\partial \mathcal{H}}{\partial p_j} = p_i \dot{q}_j = k_B T \delta^i_j.
\]  \hspace{1cm} (248)

If the Hamiltonian has a regular form, e.g.

\[
\mathcal{H} = \frac{1}{2} \sum_{i=1}^{K} A_i p_i^2 + \sum_{j=1}^{f} B_j q_j^2,
\]

with some constant \( A_i \) and \( B_j \), so that

\[
\mathcal{H} = \frac{1}{2} \sum_{i=1}^{K} p_i \frac{\partial \mathcal{H}}{\partial p_j} + \frac{1}{2} \sum_{i=1}^{f} q^j \frac{\partial \mathcal{H}}{\partial q^j},
\]

then the results (247), (248) imply that the mean\(^{105}\) system energy is related to its temperature by

\[
U \equiv \overline{\mathcal{H}} = \frac{1}{2}(K + f)k_B T,
\]  \hspace{1cm} (249)

independently of the details of the system’s microscopic dynamics. For instance this immediately implies that the heat capacity \( c_v \) per molecule of a classical monoatomic perfect gas is \( \frac{3}{2}k_B \) (\( f = 0 \), \( K = 3N \), so three “degrees of freedom” per molecule) and of a two-atomic gas \( \frac{6}{2}k_B \) (\( K = 6N \), \( f = N \), so 7 “degrees of freedom” per molecule). Similarly the heat capacity of the crystal lattice of a solid built out of \( N \) molecules and behaving

\(^{104}\)Actually even less complicated because one does not need to appeal to an artificial boundary potential.

\(^{105}\)In the case of the microcanonical ensemble representative for a macroscopically isolated system with fixed energy \( E \) (within the tolerance \( \Delta E \) which in the classical case can be set equal zero) the mean energy \( \overline{\mathcal{H}} \equiv U \) is the same as \( E \); the result is valid however also if the system is in equilibrium but does not have fixed energy.
as 3N independent harmonic oscillators\textsuperscript{106} should be \( C_V = 3Nk_B \) \((K = 3N, f = 3N;\) 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 \( \frac{1}{2}k_BT \) of energy to the mean system’s total energy, but it is clear that the implication of the general results \((247), (248)\) for the total energy depend on the precise form of the interactions.\textsuperscript{107} Needless to say, 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.

\textit{Canonical (Gibbs) Ensemble}

The Microcanonical Ensemble introduced in the preceding Lectures and applicable to 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 environment (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, that is from the point of view of thermodynamics, there should be no difference between an isolated system whose temperature (defined as \( (\partial S/\partial U)_{V,N,...} \)) 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.

To derive the Canonical Ensemble we consider the system and the heat bath (representing its environment) as a macroscopically isolated compound system and apply to it the Microcanonical Ensemble. According to the rules of the probability theory, the probability that the total energy \( E_{\text{tot}} \) of the compound system (fixed up to small uncertainty \( \Delta E \)) is distributed as \( U \) and \( E_{\text{tot}} - U \) between the system and the heat bath is given by the ratio of the number of microstates (of the compound system) corresponding to such a distribution to the total number of all microstates corresponding to the total energy \( E_{\text{tot}} \):

\[
P(U, dU) = \frac{1}{\Gamma(E_{\text{tot}}, N_{\text{h.b.}}, N_{\text{sys}}, \ldots, \Delta E)} \Delta E \omega_{\text{sys}}(U) \omega_{\text{h.b.}}(E_{\text{tot}} - U) dU, \tag{250}
\]

where

\[
\Gamma(E_{\text{tot}}, N_{\text{h.b.}}, N_{\text{sys}}, \ldots, \Delta E) = \Delta E \int_0^{E_{\text{tot}}} dU \omega_{\text{sys}}(U) \omega_{\text{h.b.}}(E_{\text{tot}} - U). \tag{251}
\]

\textsuperscript{106} 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 \( 3N \).

\textsuperscript{107} 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.
Using the Microcanonical Ensemble definition $S(E) = k_B \ln(\omega_{\text{h.b.}}(E)\Delta E)$ of the entropy of the heat bath, the probability $P(U, dU)$ can be written as

$$P(U, dU) \propto dU \omega_{\text{sys}}(U) \exp \left\{ \frac{1}{k_B} S_{\text{h.b.}}(E_{\text{tot}} - U) \right\}.$$ 

Expanding now

$$S_{\text{h.b.}}(E_{\text{tot}} - U) = S_{\text{h.b.}}(E_{\text{tot}}) - \left( \frac{\partial S_{\text{h.b.}}(E_{\text{tot}})}{\partial E_{\text{tot}}} \right)_{N_{\text{h.b.}}} U + \ldots,$$  \hspace{1cm} (252)

and neglecting all the terms of higher order in $U$ than the first one on account of the fact that in the limit of the very large heat bath the probability $P(U, dU)$ of distributions of energy markedly different than ones in which $U/E_{\text{tot}} \sim N_{\text{sys}}/N_{\text{h.b.}}$ is negligible, while $S_{\text{h.b.}} \sim N_{\text{h.b.}}$ and $E_{\text{tot}} \sim N_{\text{h.b.}}$ (the $n$-th derivative of $S_{\text{h.b.}}$ with respect to energy is then $\sim N_{\text{h.b.}}^{-n}$ and the terms of higher order in $U$ than the first one are suppressed with respect to it by powers of $N_{\text{sys}}/N_{\text{h.b.}}$), one can (in the limit of infinitely large heat bath) write the Canonical Ensemble distribution $ho_U(U)$ of the system energy in the form

$$\rho_U(U) = \text{Const.} \omega_{\text{sys}}(U) e^{-U/k_B T} \equiv \text{Const.} \omega_{\text{sys}}(U) e^{-\beta U}.$$  \hspace{1cm} (253)

We have introduced here the traditional symbol

$$\beta \equiv \frac{1}{k_B T}.$$  \hspace{1cm} (254)

The normalization constant is of course given by$^{108}$

$$\text{Const.}^{-1} = \int_0^\infty dU \omega_{\text{sys}}(U) e^{-U/k_B T}.$$ 

The Canonical Ensemble energy distribution (253) 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 a heat bath at temperature $T$ can be obtained by marginalization (in the language of statistics) of the joint distribution function $\rho(q, p, q_{\text{h.b.}}, p_{\text{h.b.}})$ of the microcanonical ensemble representative for the system and the heat bath as a single compound, macroscopically isolated system:

$$\rho(q, p) = \frac{1}{\Gamma(E_{\text{tot}}, N_{\text{h.b.}}, N_{\text{sys}}), \ldots, \Delta E)} \int d\Gamma_{(q_{\text{h.b.}}, p_{\text{h.b.}})} \rho(q, p, q_{\text{h.b.}}, p_{\text{h.b.}})$$

$$\propto \int_{E_{\text{tot}} - \mathcal{H}_{\text{sys}} \leq \mathcal{H}_{\text{h.b.}} \leq E_{\text{tot}} + \Delta E - \mathcal{H}_{\text{sys}}} d\Gamma_{(q_{\text{h.b.}}, p_{\text{h.b.}})} = \exp \left\{ \frac{1}{k_B} S_{\text{h.b.}}(E_{\text{tot}} - \mathcal{H}_{\text{sys}}(q, p)) \right\}.$$

$^{108}$To make it clear: the probability distribution defined by (250) and (251) is normalized by construction. However, once the expansion (252) has been truncated, the tails of the distribution are modified (insignificantly from the practical point of view) and its normalization must be readjusted.
It has been assumed here that the Hamiltonian of the compound system is the sum \( \mathcal{H}_{\text{sys}} + \mathcal{H}_{\text{h.b.}} \); in other words its part \( \mathcal{H}_{\text{sys-h.b.}} \) - the interaction of the system with the heat bath ensuring in fact the assumed thermal contact between them - has been assumed to be negligible.\(^{109}\) Expanding \( S_{\text{h.b.}}(E_{\text{tot}} - \mathcal{H}_{\text{sys}}(q,p)) \) in the power series in \( \mathcal{H}_{\text{sys}}(q,p) \), rejecting terms of the expansion of order higher than the first one (again, on account of the fact that they affect the constructed distribution function \( \rho(q,p) \) vanishingly 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 \( Z_{\text{stat}} \), one obtains the distribution function \( \rho(q,p) \) of the Canonical Ensemble in the form

\[
\rho(q,p) = \frac{1}{Z_{\text{stat}}} \exp\{-\mathcal{H}_{\text{sys}}(q,p)/k_B T\} .
\]

(255)

Its normalization factor

\[
Z_{\text{stat}}(T, V, \ldots, N) = \int d\Gamma(q,p) \exp\{-\mathcal{H}_{\text{sys}}(q,p)/k_B T\} .
\]

(256)

which is a function of the temperature \( T \), the system’s volume \( V \), the number \( N \) (numbers \( N_1, \ldots, N_r \) of particles constituting it and possibly other macroscopic parameters (denoted by ellipses) characterizing the system is called the Canonical Statistical Sum or the Canonical Partition Function. By introducing under the integral in (256) the unity written as \( 1 = \int dE \delta(E - \mathcal{H}(q,p)) \) and exchanging the order of integrations, the canonical partition function can also be written in the alternative form as

\[
Z_{\text{stat}}(T, V, \ldots, N) = \int_0^\infty dE \omega(E) \exp\{-E/k_B T\} .
\]

(257)

As will be shown below, it contains the complete thermodynamic information about the system.

The derivation of the statistical operator \( \hat{\rho} \) of the quantum Canonical Ensemble representative for the system in thermal equilibrium with the heat bath is as follows. The basis of the Hilbert space of the entire isolated compound system \( \mathcal{H} = \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{h.b.}} \) is formed by the state vectors of the form \( |l\rangle \otimes |l_{\text{h.b.}}\rangle \) which are assumed to be eigenvectors of the Hamiltonian \( \mathcal{H}_{\text{sys}} \otimes \hat{1}_{\text{h.b.}} + \hat{1}_{\text{sys}} \otimes \mathcal{H}_{\text{h.b.}} \); the interaction term \( \mathcal{H}_{\text{h.b.-sys}} \) ensuring the thermal contact between the system and the heat bath is, as in the classical case and with the same justification, assumed to be negligible. 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 \( E_{\text{tot}} \) to \( E \))

\[
\hat{\rho}_{\text{Micro}} = \frac{1}{\Gamma(E \ldots, \Delta E)} \sum_{l_{\text{h.b.}}, l} (|l\rangle \otimes |l_{\text{h.b.}}\rangle)(\langle l_{\text{h.b.}}| \otimes \langle l|)
\]

\[
E \leq E_{\text{h.b.}} + E_l \leq E + \Delta E
\]

\(^{109}\)In fact, since the role of the heat bath is auxiliary only, this can freely be assumed. In connection with this it should be recalled that the systems forming the ensemble are isolated - only the real system, of which the systems of the ensemble are supposed to be representative, is in contact with the heat bath but it is assumed to have reached the equilibrium with the heat bath before it starts to be represented by the ensemble; so reaching this equilibrium might have taken arbitrarily long time.
$$\hat{\rho} = \frac{1}{\Gamma(E \ldots, \Delta E)} \sum_l |l\rangle \langle l| \otimes \sum_{l_{h.b.}} |l_{h.b.}\rangle \langle l_{h.b.}|$$

$$E_l \leq E + \Delta E \quad E - E_l \leq E_{h.b.} \leq E + \Delta E_{\text{tot}} - E_l$$

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 \(\mathcal{H}_{h.b.}\) of the heat bath. The sum over \(l_{h.b.}\) 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 \(\Gamma(E - E_l, N_{h.h., \ldots, \Delta E}) = \exp(S_{h.b.}(E - E_l)/k_B)\). This, upon expanding up to first order in \(E_l\) and including the zeroth order term in the normalization, leads to the final result

$$\hat{\rho} = \frac{1}{Z_{\text{stat}}} \sum_l |l\rangle \langle l| \exp(-E_l/k_B T) \equiv \frac{1}{Z_{\text{stat}}} \exp\left(-\mathcal{H}/k_B T\right), \quad (258)$$

with the normalization factor (the quantum Canonical Ensemble partition function) \(Z_{\text{stat}}\) given by

$$Z_{\text{stat}} = \text{Tr} \left( e^{-\mathcal{H}/k_B T} \right) = \sum_l \langle l| e^{-\mathcal{H}/k_B T} |l\rangle = \sum_{E_n} \exp(-E_l/k_B T) = \sum_{E_n} d_n \exp(-E_n/k_B T). \quad (259)$$

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 \(d_n\). This second form of the formula is the quantum case counterpart of the classical formula (257). It should be noted that both results (255) in the classical case and (258) have the forms (196) and (230), respectively, established by considering general properties of macroscopic systems.

We will now show that if the system is simple in the thermodynamical sense, that is only volume work can be done on it, the statistical sum given by (256) or (257) in the classical case and by (259) in the quantum case, is directly related to the free Helmholtz energy \(F\) which is already known to be the proper thermodynamic potential characterizing a system kept at constant temperature by thermal contact with a heat bath. This readily follows from the “golden formulae” (246) or (245). Indeed, considering for illustration the classical case, (246) applied to (255) yields:

$$S = -k_B \ln \rho = -k_B \left(-\mathcal{H}/k_B T - \ln Z_{\text{stat}}\right) = \frac{1}{T} \mathcal{H} + k_B \ln Z_{\text{stat}}. \quad (258)$$

\(^{110}\)More precisely, its matrix elements \(|l'\rangle \langle l|\) are identified with

$$\sum_{l_{h.b.}} \langle l_{h.b.}| \otimes \langle l'| \hat{\rho}_\text{Micro} |l\rangle \otimes |l_{h.b.}\rangle.$$

\(^{111}\)Justification of the rejection of all higher order terms of the expansion is the same as in the classical case.

\(^{112}\)Notice also that the symbol \(\beta\) in the forms (196) and (230) has now acquired the interpretation (254).
After a small rearrangement and identifying \( \mathcal{H} \) with the system’s internal energy \( U \), this means that (recall, that \( F = U - TS \))

\[
F(T, V, N) = -k_B T \ln Z_{\text{stat}}(T, V, N).
\] (260)

Obviously this formula stays valid in the quantum case (the steps are in that case completely analogous).

The rule (260) is also valid if the system consists of more than one material component - simply \( N \) should be replaced by the numbers \( N_1, \ldots, N_r \) of different kinds of particles. 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 Canonical Ensemble formalism applied to a mixture of different classical perfect gases readily leads to the Gibbs Ansatz (158) and, as discussed in Lectures VIII and XI, it is the product of factors \( N_1! \ldots N_r! \) which leads to the mixing entropy.

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, mean value of an observable \( O \) (represented by a phase space function \( O(q,p) \) in the classical case and by a Hermitian operator \( \hat{O} \) in the quantum case) can be computed as

\[
\overline{O} = \int d\Gamma_{(q,p)} O(q,p) \rho(q,p) = \frac{1}{Z_{\text{stat}}} \int d\Gamma_{(q,p)} O(q,p) e^{-\mathcal{H}(q,p)/k_B T},
\]

\[
\overline{O} = \text{Tr}(\hat{O} \hat{\rho}) = \frac{1}{Z_{\text{stat}}} \text{Tr} \left( \hat{O} e^{-\mathcal{H}/k_B T} \right),
\] (261)

in the classical and quantum cases, respectively. Similarly can be computed also its mean quadratic fluctuation (190) to estimate its relative fluctuation (191). Since 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 \( U \) and its fluctuation can be obtained in general terms, without using the explicit forms of these Hamiltonians. Indeed, from the formulae (261) it follows that (for definiteness we consider the quantum case - classical calculation is completely analogous)

\[
U = \overline{E} = \frac{1}{Z_{\text{stat}}} \text{Tr} \left( \mathcal{H} e^{-\mathcal{H}/k_B T} \right) = -\frac{1}{Z_{\text{stat}}} \frac{\partial}{\partial \beta} \text{Tr} \left( e^{-\beta \mathcal{H}} \right) \bigg|_{\beta=1/k_B T} = -\frac{\partial}{\partial \beta} \ln Z_{\text{stat}} \bigg|_{\beta=1/k_B T}.
\] (262)
Using the formula (260) and noticing that $\partial/\partial \beta = -k_B T^2 \partial/\partial T$ 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 (146). The mean quadratic fluctuation of $U$ can be expressed in the similarly general way. To this end one starts with (we suppress the instruction to set $\beta = 1/k_B T$ at the end)

$$\overline{E^2} = \frac{1}{Z_{\text{stat}}} \text{Tr} \left( \hat{\mathcal{H}}^2 e^{-\hat{\mathcal{H}}/k_B T} \right) = -\frac{1}{Z_{\text{stat}}} \frac{\partial}{\partial \beta} \text{Tr} \left( \hat{\mathcal{H}} e^{-\hat{\mathcal{H}}/k_B T} \right),$$

$$= -\frac{1}{Z_{\text{stat}}} \frac{\partial}{\partial \beta} \left( Z_{\text{stat}} \overline{E} \right) = -\frac{\partial \overline{E}}{\partial \beta} \overline{E} - \overline{E} \frac{\partial \ln Z_{\text{stat}}}{\partial \beta} \equiv -\frac{\partial \overline{E}}{\partial \beta} + \overline{E^2}.$$

It follows that

$$\sigma^2_U \equiv \overline{E^2} - \overline{E}^2 = -\frac{\partial \overline{E}}{\partial \beta} = k_B T^2 \left( \frac{\partial U}{\partial T} \right)_{V,N} = k_B T^2 C_V. \quad (263)$$

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, $C_V = N c_v$ and $U = Nu$, the relative fluctuation of its energy is suppressed by $1/\sqrt{N}$

$$\frac{\sqrt{\sigma^2_U}}{U} = \frac{1}{Nu} \sqrt{N k_B T^2 c_V} \sim 1/\sqrt{N},$$

in agreement with the general result (194).

It is also worth to ask what thermodynamical potential is related to the Canonical Ensemble statistical sum $Z_{\text{stat}}$ 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”)

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_{\text{int}} - \sum_{i=1}^N \hat{\mathbf{\mu}}_i \cdot \mathbf{\mathcal{H}},$$

where the second term is the coupling of magnetic moments $\mathbf{\mu}_i$ of the system’s individual elements (molecules, for instance) to an external magnetic field $\mathbf{\mathcal{H}}$ through the operators $\hat{\mathbf{\mu}}_i$ representing them.\footnote{In such a setting there is no distinction between the magnetic field $\mathbf{\mathcal{H}}_0$ produced by the experimental setup - by a current passing through a coil, for instance - in the absence of the magnetic material and the actual magnetic field $\mathbf{\mathcal{H}}_0$ when the magnetic material is present, because the magnetic moments $\mathbf{\mu}$ 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 $-\sum_{i \neq j} J_{ij} \hat{\mathbf{\mu}}_i \cdot \hat{\mathbf{\mu}}_j$ (which, if present, are here included in $\hat{\mathcal{H}}_{\text{int}}$).}

Assuming such a system to be in equilibrium with a heat bath
at temperature $T$ and applying the “golden formula” (245) to the Canonical Ensemble representing it we get

$$S = -k_B \left( -\frac{1}{k_BT} \left( E_{\text{int}} - \sum_i \mu_i \cdot \mathcal{H} \right) \right) + k_B \ln Z_{\text{stat}}$$

$$= \frac{1}{T} E_{\text{int}} - \frac{1}{T} \mathcal{H} \cdot \sum_i \mu_i + k_B \ln Z_{\text{stat}},$$

Identifying now $E_{\text{int}}$ with the system’s internal energy $U$ and $\sum_i \mu_i$ with the system’s total magnetization $\mathbf{M}$, one obtains

$$-k_B T \ln Z_{\text{stat}}(T, V, \mathcal{H}, N) = U - TS - \mathbf{M} \cdot \mathcal{H},$$

which is the magnetic Gibbs function $G(T, V, \mathcal{H}, N)$ whose differential is\textsuperscript{114}

$$dG = -SdT - p dV - \mathbf{M} \cdot d\mathcal{H} + \mu dN.$$  

The differential of the internal energy is then as usually (in this Course)

$$dU = T dS - p dV + \mathcal{H} \cdot d\mathbf{M} + \mu dN.$$  

The reader should be warned, however, that there is another school of authors which include the interaction of magnetic moments into what they call the system’s internal energy (let denote it $\tilde{U}$); the statistical sum is then related to the function $F$

$$-k_B T \ln Z_{\text{stat}}(T, V, \mathcal{H}, N) = F = \tilde{U} - TS,$$  

This is however not a mere change of the notation (using $F$ in place of $G$): since

$$\tilde{U} = U - \mathbf{M} \cdot \mathcal{H},$$  

the differential of the “internal energy” $\tilde{U}$ is

$$d\tilde{U} = T dS - p dV - \mathbf{M} \cdot d\mathcal{H} + \mu dN.$$  

so that the elementary work of magnetization is now $-\mathbf{M} \cdot d\mathcal{H}$ (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

$$Z_{\text{stat}} = \text{Tr} \left( e^{-\mathcal{H} / k_BT} \right), \quad \overline{O} = \frac{1}{Z_{\text{stat}}} \text{Tr} \left( \hat{O} e^{-\mathcal{H} / k_BT} \right),$$

(264)

\textsuperscript{114}The chemical potential $\mu$ should not be confused with the magnetic moment $\mathbf{\mu}$. It is clear that the SI system of units which brings in yet another $\mu_0$ (the magnetic susceptibility of vacuum) is utterly inconvenient here!
of the quantum Canonical Ensemble requires specifying the Hilbert space of the system of interest in which act the operators \( \hat{H}, \hat{\rho}, \hat{O} \). If the system is, as usually is the case, composed 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 particles.\(^{115}\) The Hilbert space of the system can be then constructed as a tensor product of single-particle Hilbert spaces spanned by state-vectors representing states of individual particles. That is, one first specifies a basis of state-vectors representing quantum states of a single particle taking for instance generalized eigenvectors \(|\mathbf{x}\rangle\) of the position operator,\(^{116}\) or eigenvectors \(|\mathbf{p}\rangle\) of the momentum operator, or if the particle has a nonzero spin the vectors \(|\mathbf{p}, \sigma\rangle\) where \( \sigma \) 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 couple of independent labels). The vectors \(|l_i\rangle\) thus span the single-particle Hilbert space \( \mathcal{H}_i \) of the \( i \)-th particle of the system. The Hilbert space \( \mathcal{H}_{(N)} \) of the entire system is then constructed as \( \mathcal{H}_{(N)} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_N \), that is, it spanned by the basis state vectors of the form

\[
|l_1\rangle \otimes \ldots \otimes |l_N\rangle. 
\]  

Most of physical systems however are composed of identical indistinguishable particles (or several groups of identical particles 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 in the course of its development were abstracted from the experimental facts dictate that the physical states of such systems are represented not by arbitrary superpositions of state-vectors (265) but only 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; in two space dimensions other symmetry properties are also possible and the corresponding particles are called anyons. Therefore the Hilbert spaces of a system of \( N \) fermions or of \( N \) bosons are spanned by the following basis vectors:

\[
\frac{1}{\sqrt{N!}} \sum_P (-1)^P |l_{P(1)}\rangle \otimes \ldots \otimes |l_{P(N)}\rangle \equiv |l_1, l_2, \ldots, l_N\rangle \quad \text{fermions,}
\]

\(^{115}\)One can sometimes meet the statements to the effect that the construction of the Hilbert space of the 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 physically decide what are the possible quantum states (in the abstract sense) of the system - what states of its individual elements can be physically identified - and only then model them by the choice of the appropriate Hilbert space in which the action of operators is realized.

\(^{116}\)These better be avoided as the position operator does not exist in relativistic physics!
$$\frac{1}{\sqrt{N!}} \sum_P |l_{P(1)}\rangle \otimes \cdots \otimes |l_{P(N)}\rangle \equiv |l_1, l_2, \ldots, l_N\rangle \text{ bosons.} \quad (266)$$

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 symmetry of the vectors (266)) 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 be in this case 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 (266) are automatically normalized to unity, while in the case of they require additional normalization if some labels $l_i$ are equal.\footnote{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 $1/\sqrt{n_l!}$.}

As the Hamiltonian $\mathcal{H}_N$ of the system of $N$ particles is supposed to be built (as appropriate tensor products) of operators acting in single-particle Hilbert spaces of individual particles, the traces in the formulae (264) can be computed using the bases (266):

$$\text{Tr}(e^{-\mathcal{H}/k_B T}) = \frac{1}{N!} \sum_{l_1=1}^{\infty} \cdots \sum_{l_N=1}^{\infty} \langle l_N, \ldots, l_1|e^{-\mathcal{H}/k_B T}|l_1, \ldots, l_N\rangle,$$

and

$$\text{Tr}(\hat{\rho} \hat{O}) = \frac{1}{N!} \sum_{l_1=1}^{\infty} \cdots \sum_{l_N=1}^{\infty} \frac{1}{N!} \sum_{l'_1=1}^{\infty} \cdots \sum_{l'_N=1}^{\infty} \langle l_N, \ldots, l_1|\hat{\rho}|l'_1, \ldots, l'_N\rangle \langle l'_N, \ldots, l'_1|\hat{O}|l_1, \ldots, l_N\rangle.$$

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 $l_i$.

It should be stressed that in this form the formulae are completely general and include all possible quantum effects. In particular of the mythical “quantum statistics”.

Yet the adopted notation $|l_1, \ldots, l_N\rangle$ of the basis state-vectors is rather inconvenient. It is much more practical to pass to the so called occupation number representation in which the numbers in kets tell how many particles occupies successive single-particle states. Thus we set

$$|n_1, n_2, \ldots\rangle = \frac{1}{\sqrt{n_1!n_2!\ldots}}|1, \ldots, 1, 2, \ldots, 2, \ldots\rangle.$$

The square root of factorials makes these vectors well normalized. Of course in the case of fermions $n_i = 0$ or 1 only. Moreover, the sum of the occupation numbers $n_i$ must always
be $N$. For example, the (normalized to unity) vector $|1, 1, 1, 2, 3, 3, 7, 7, 11, 13\rangle/\sqrt{3!2!2!}$ of 10 bosons is in the occupation number representation written as

$$|3, 1, 2, 0, 0, 0, 2, 0, 0, 0, 1, 0, 1, 0, 0, 0, \ldots \rangle.$$

It should be stressed that this is not a change of the basis in the Hilbert space but only a change of notation. It is also to important to note that in the notation $|l_1, \ldots, l_N\rangle$ the number of entries in the ket is finite, but each label $l_i$ can assume infinitely many (discrete values); in the notation $|n_1, n_2, \ldots\rangle$, the number of entries in the ket is infinite but the values of the labels $n_i$ are restricted by the condition $n_1 + n_2 + \ldots = N$ (and, moreover, if particles are fermions each $n_i$ is either zero or one). The occupation number notation of state-vectors of $N$ bosons should be also contrasted (in order to avoid confusion) with the basis state-vectors $|n_1, \ldots, n_N\rangle$ of a system of $N$ quantum harmonic oscillators: in this case it is the number of the labels $n_i$ which is $N$ but each $n_i$ can run from 0 to infinity. We will see that if the restriction $n_1 + n_2 + \ldots = N$ is removed - and we will remove it passin
to the Grand Canonical Ensemble - the system of bosons will become mathematically identical with a system of (infinitely many) harmonic oscillators.

The usefulness of the introduced notation will become evident when we introduce the creation and annihilation operators (associated with single-particle states) through which all operators of interest acting in the system's Hilbert space - the Hamiltonian $\hat{H}$ (including its interaction terms) of the system as well as all observables - can be expressed. Here we will consider only a system of $N$ mutually noninteracting identical molecules whose Hamiltonian has the general form

$$\hat{H}_N = \sum_{i=1}^{N} \hat{1} \otimes \ldots \otimes \hat{H}(i) \otimes \ldots \otimes \hat{1},$$

($\hat{H}(i)$ at the $i$-th position) of a sum of $N$ operators each of which acts essentially only in one single-particle Hilbert space constructed as a (anti)symmetrized product of $N$ (isomorphic) single-particle Hilbert spaces. If the vectors $|l\rangle$ are chosen to be eigenvectors of the single particle Hamiltonian $\hat{H}$ with with the eigenvalues $\varepsilon_i$, then the action of $\hat{H}_N$ on the basis vectors $|n_1, n_2, n_3, \ldots\rangle$ of the Hilbert space of $N$ molecules is particularly simple

$$\hat{H}_N |n_1, n_2, n_3, \ldots\rangle = \left( \sum_{i} n_i \varepsilon_i \right) |n_1, n_2, n_3, \ldots\rangle.$$

The statistical sum $Z_{\text{stat}}$ of such a system of $N$ mutually noninteracting molecules is given by the expression

$$Z_{\text{stat}} = \sum_{n_1=0}^{n_{\text{max}}} \sum_{n_2=0}^{n_{\text{max}}} \ldots \delta_{N, \sum_i n_i} e^{-n_1 \varepsilon_1/k_B T} e^{-n_2 \varepsilon_2/k_B T} \ldots, \quad (267)$$

in which $n_{\text{max}} = 1$, if the molecules are fermions and $n_{\text{max}} = \infty$, if they are bosons. Unfortunately, even in this simple case the statistical sum cannot be computed easily
(except for special forms of the spectra $\varepsilon_l$ of the Hamiltonian of a single molecule) because of the presence of the Kronecker delta which expresses the condition of constancy of molecules in the Canonical Ensemble. There are two ways of going around this difficulty: one is to use the standard trick of statistical physics (and thermodynamics): in order to control the quantity which is conserved 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 it with the heat bath at fixed temperature and by controlling 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 is representative for the system exchanging matter (molecules) with a reservoir at fixed chemical potential $\mu$. Another way is to use the so called Boltzmann approximation which we now discuss.

In the Boltzmann approximation the expression (267) is replaced by

$$Z_{\text{stat}} = \frac{1}{N!} \left( \sum_l e^{-\varepsilon_l/k_\beta T} \right)^N. \quad (268)$$

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 (268) 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 of the level of energy $n_1\varepsilon_1 + n_2\varepsilon_2 + \ldots$ the factor

$$\binom{N}{n_1} \binom{N - n_1}{n_2} \binom{N - n_1 - n_2}{n_3} \ldots = N! \left[ \prod_{a=1}^{\infty} (n_a!) \right]^{-1}.$$

Notice that this counting treats particles as distinguishable (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 (265) and not only by its totally symmetrized or totally antisymmetrized subspaces. Indistinguishability of particles is here only taken into account by dividing all these factors by $N!$. 

166
With this counting the sum (267) is replaced by

$$Z_{\text{stat}} = \frac{1}{N!} \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \cdots \frac{N!}{\prod_{a=1}^{\infty} (n_a!)} \delta_{N, \sum_i n_i} e^{-n_1 \varepsilon_1 / k_B T} e^{-n_2 \varepsilon_2 / k_B T} \ldots,$$

(269)

This precisely gives\(^{118}\) the expression (268).

It is instructive to see on a simple example, what is the difference between the sums (267) in the cases when particles are bosons or fermions and the sum in (268). Let’s take \(N = 3\) particles and introduce the notation \(x_i = \exp(-\varepsilon_i / k_B T)\). If \(N = 3\) the formula (267) yields

\[
Z_{\text{stat}} = x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_1 x_2 x_3 + \ldots \quad \text{bosons,}
\]

\[
Z_{\text{stat}} = x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 + \ldots \quad \text{fermions,}
\]

while from the formula (268) we obtain (in both cases)

\[
Z_{\text{stat}} = \frac{1}{3!} (x_1 + x_2 + x_3 + x_4 + \ldots)^3 = \frac{1}{6} (x_1^3 + 3x_1^2 x_2 + 3x_1 x_2^2 + 6x_1 x_2 x_3 + \ldots).
\]

It is clear that contributions to the statistical sum \(Z_{\text{stat}}\) of singly occupied single-particle states in (267) are properly accounted in the approximation (268), but those of multiply occupied single-particle states are not: if particles are fermions they are totally absent in (267) whereas if particles are bosons they have higher weights.

It follows that the Boltzmann approximation (268) 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 \(\mu\). 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

\[
\overline{n}_l = \frac{1}{1 + \exp((\varepsilon_l - \mu) / k_B T)} \quad \left\{ \begin{array}{ll}
\text{bosons} & \\
\text{fermions} &
\end{array} \right.
\]

This is smaller than 1 (that is, the probability of the occupation of a given energy state \(|l\rangle\) is low) independently of the number \(l\) 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:

\[
z \equiv e^{\mu / k_B T} \ll 1.
\]

\(^{118}\)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 \(((x_1 + x_2 + x_3 + \ldots)^N)\).
This is typical 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:

\[ \mu = -k_B T \ln \left[ \frac{V}{N} \left( \frac{mk_B T}{2\pi \hbar^2} \right)^{3/2} \right]. \]

It becomes large negative at high temperatures (and moreover, \( \mu/k_B T \to -\infty \) as \( T \to \infty \)). More precisely, the condition (270) translates into the condition

\[ \frac{N}{V} \left( \frac{mk_B T}{2\pi \hbar^2} \right)^{3/2} \gg 1, \]

or

\[ \lambda_T \equiv \left( \frac{2\pi \hbar^2}{mk_B T} \right)^{1/2} \ll \left( \frac{V}{N} \right)^{1/3}, \]

which means that the so-called thermal wavelength \( \lambda_T \) of molecules should be small compared to the mean intermolecular distance.
LECTURE XIII (STAT)

Grand Canonical Ensemble
As we have 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 the Hilbert space of states of identical elements of the system: even if these identical elements (particles) are mutually noninteracting, there is an intrinsic quantum entanglement of their states resulting from the requirements of symmetry. As a result the canonical statistical sum $Z_{\text{stat}}$ cannot be computed (except for special cases). Technically, it is the condition of constant number of particles (elements) which prevents computation of $Z_{\text{stat}}$. Therefore it is convenient to consider most of the quantum problems within the Grand Canonical Ensemble of systems which formally is the ensemble representative for a system in equilibrium with a very big reservoir at temperature $T$ and chemical potential $\mu$ with which the system of interest exchanges energy and matter (particles). Although the number of particles is then variable (different isolated and closed systems of the ensemble have different internal energies and different numbers of particles), in all situations in which fluctuations of the number of particles around the mean $\overline{N}$ value are negligible (essentially proportional to $1/\sqrt{N}$), the results obtained using this this ensemble are from the practical point of view the same as the ones that would be obtained using the Canonical Ensemble.

We begin by deriv ing the distribution of energy and the number of particles in the system given by this ensemble - a quantity which os independent of wheter the system is treated classically of quantum mechanically. Thus we assume that the real supersystem consists of a big reservoir weakly interacting with the system of interest and that particles (or molecules) which are indestructible can be exchanged between the system and the reservoir. The entire supersystem is macroscopically isolated and characterized by the total energy $E_{\text{tot}}$ (with some uncertainty $\Delta E$) and the total number of particles $N_{\text{tot}}$. The corresponding microcanonical ensemble representative of the real physical system consists of a number $N$ ($N \to \infty$ in the limit) of supersystems composed of isolated from one another reservoir and the system; in elements of the ensemble the total energy $E_{\text{tot}}$ and the total number of particles $N_{\text{tot}}$ are distributed between the system and the reservoir in all possible ways and - in agreement with the principle of a priori equal probabilities - each microstate of the supersystem characterized by a concrete distribution of $E_{\text{tot}}$ and $N_{\text{tot}}$ and some microstates of the system and of the reservoir corresponding to this distribution of $E_{\text{tot}}$ and $N_{\text{tot}}$ has the same statistical weight. According to the rules of the probability theory the probability $\rho_V(N, U)dU$ that there are $N$ particles in the system (and $N_{\text{tot}} - N$ particles in the reservoir) and its energy is between $U$ and $dU$ is given by the ratio of the number of microstates (the number of supersystems in the ensemble) realizing this situation to the total number of all microstates (the subscripts “res” and “sys” refer to the reservoir and the system, respectively)

$$\rho_V(N, U) dU = \text{Const. } \omega_{\text{sys}}(U, N) \omega_{\text{res}}(E_{\text{tot}} - U, N_{\text{tot}} - N) \Delta E dU,$$
where

\[
\text{Const.}^{-1} = \sum_{N=0}^{N_{\text{tot}}} \int_{0}^{E_{\text{tot}}} dU \, \omega_{\text{res}}(E_{\text{tot}} - U, N_{\text{tot}} - N) \Delta E.
\]

The quantity \( \omega_{\text{res}}(E_{\text{tot}} - U, N_{\text{tot}} - N) \Delta E \) is however directly related to the entropy of the reservoir treated as a macroscopically isolated system with energy \( E_{\text{tot}} - U \) and the number \( N_{\text{tot}} - N \) of particles. Thus

\[
\rho_{U}(N, U) \propto dU \, \omega_{\text{sys}}(U, N) \exp \left\{ \frac{1}{k_{B}} S_{\text{res}}(E_{\text{tot}} - U, N_{\text{tot}} - N) \right\},
\]

and, upon expanding this entropy to the first order in \( U \) and \( N \) (the justification cutting the expansion is essentially the same as in the case of the Canonical Ensemble) and normalizing the distribution anew (to correct for the inessential change in the tails of the distribution introduced by dropping higher order terms of the expansion), one obtains the distribution\(^{119}\)

\[
\rho_{U}(U, N) = \text{Const.} \, \omega_{\text{sys}}(U, N) \exp \left\{ -\frac{1}{k_{B} T} U + \frac{\mu}{k_{B} T} N \right\}, \quad (271)
\]

in which \( T \) and \( \mu \) are to be identified with the temperature and the chemical potential of the reservoir and

\[
\text{Const.}^{-1} = \sum_{N=0}^{\infty} \int_{0}^{\infty} dU \, \omega_{\text{sys}}(U, N) \exp \left\{ -\frac{1}{k_{B} T} U + \frac{\mu}{k_{B} T} N \right\},
\]

The corresponding classical distribution \( \rho(q, p, N) \) function of the Grand Canonical Ensemble pertaining to the system alone is derived by marginalizing (in the statistical sense) the distribution of the supersystem microcanonical ensemble

\[
\rho_{\text{micro}} = \begin{cases} 
\text{Const.}, & \text{if } E_{\text{tot}} \leq \mathcal{H}_{\text{sys}}^{(N)} + \mathcal{H}_{\text{res}}^{(N_{\text{tot}} - N)} \leq E_{\text{tot}} + \Delta E \\
0, & \text{otherwise}
\end{cases}
\]

in which \( \text{Const.}^{-1} = \Gamma(E_{\text{tot}}, V_{\text{res}}, V_{\text{sys}}, N_{\text{tot}}, \Delta E) \) is given by

\[
\sum_{N=0}^{N_{\text{tot}}} \int d\Gamma_{\text{sys}}^{(N)} \int d\Gamma_{\text{res}}^{(N_{\text{tot}} - N)} \theta(E_{\text{tot}} + \Delta E - \mathcal{H}_{\text{sys}}^{(N)} - \mathcal{H}_{\text{res}}^{(N_{\text{tot}} - N)}) \theta(\mathcal{H}_{\text{sys}}^{(N)} + \mathcal{H}_{\text{res}}^{(N_{\text{tot}} - N)} - E_{\text{tot}}),
\]

with respect to the variables of the reservoir. That is,

\[
\rho(q, p, N) = \text{Const.} \int_{E_{\text{tot}} - \mathcal{H}_{\text{sys}}^{(N)} \leq \mathcal{H}_{\text{res}}^{(N_{\text{tot}} - N)} \leq E_{\text{tot}} - \mathcal{H}_{\text{sys}}^{(N)} + \Delta E} d\Gamma_{\text{res}}^{(N_{\text{tot}} - N)}
\]

\[
= \text{Const.} \Gamma_{\text{res}}(E_{\text{tot}} - \mathcal{H}_{\text{sys}}^{(N)}; V_{\text{res}}, N_{\text{tot}} - N).
\]

\(^{119}\)One simultaneously takes the limits \( E_{\text{tot}} \to \infty, N_{\text{tot}} \to \infty \) with the ratio \( E_{\text{tot}}/N_{\text{tot}} \) kept constant; in this way the distribution extends to an arbitrarily big number \( N \) of particles and arbitrarily large energy \( U \) of the system.
This is again proportional to the exponential of the reservoir’s entropy \((1/\kappa_B)S_{\text{res}}(E_{\text{tot}} - \mathcal{H}_{\text{sys}}(N), V_{\text{res}}, N_{\text{tot}} - N)\), so upon expanding to the first order and taking the limits \(E_{\text{tot}} \to \infty, N_{\text{tot}} \to \infty\) with the ratio \(E_{\text{tot}}/N_{\text{tot}}\) kept constant one arrives at the final result

\[
\rho(q, p, N) = \frac{1}{\Xi_{\text{stat}}} \exp \left\{ -\frac{1}{k_B T} \left( \mathcal{H}^{(N)}(q, p) - \mu N \right) \right\},
\]

in which the Grand Canonical Ensemble statistical sum \(\Xi_{\text{stat}}(T, V, \mu)\) (in case only volume work can be done on the system in question) is given by

\[
\Xi_{\text{stat}} = \sum_{N=0}^{\infty} \int d\Gamma^{(N)} \exp \left\{ -\frac{1}{k_B T} \left( \mathcal{H}^{(N)}(q, p) - \mu N \right) \right\}
\]

\[
= \sum_{N=0}^{\infty} e^{\mu N/k_B T} \int d\Gamma^{(N)} \exp \left( -\mathcal{H}^{(N)}(q, p)/k_B T \right) \equiv \sum_{N=0}^{\infty} e^{\mu N/k_B T} Z_{\text{stat}}(T, V, N). \tag{273}
\]

We can now use the entropy golden formula to relate the statistical sum \(\Xi_{\text{stat}}\) to the appropriate thermodynamical potential:

\[
S = -k_B \bar{\ln} \rho = -k_B \sum_{N=0}^{\infty} \int d\Gamma^{(N)} \rho(q, p, N) \ln \rho(q, p, N)
\]

\[
\quad = \frac{1}{T} \sum_{N=0}^{\infty} \int d\Gamma^{(N)} \rho(q, p, N) \left( \mathcal{H}^{(N)}(q, p) - \mu N \right) + k_B \Xi_{\text{stat}}.
\]

Identifying the first term with \(\bar{E}/T\), the second one with \(\mu \bar{N}/T\) and rearranging the formula one obtains (we denote \(\bar{E}\) by \(U\) and drop the bar on \(N\))

\[
-k_B T \ln \Xi_{\text{stat}}(T, V, \mu) = U - TS - \mu N. \tag{274}
\]

Thus, the Grand Canonical Ensemble partition function is related to the thermodynamical potential (144)

\[
\Omega(T, V, \mu) = U - TS - \mu N - V p(T, \mu). \tag{275}
\]

(The last form follows from extensiveness - see Lectures VII and VIII.) It is given as a function of its natural variables and, therefore, conveys full thermodynamic information about the system. 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 \(\mathcal{H}\), the function \(-k_B T \ln \Xi_{\text{stat}}(T, V, \mathcal{H}, \mu)\) is the appropriate thermodynamical potential \(\Phi = U - TS - \mathcal{H} \cdot \mathbf{M} - \mu N\) (provided we continue not to include the interaction of the system magnetization with the external field into internal energy).

Before we embark ourselves in discussing 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) as given by the classical version of the ensemble.
To compute the mean quadratic fluctuation $\sigma_N^2 = (N - \bar{N})^2 = \bar{N}^2 - \bar{N}^2$ one uses the standard trick and writes ($\beta \equiv 1/k_B T$)

$$\bar{N}^2 = \frac{1}{\Xi_{\text{stat}}} \sum_{N=1}^{\infty} N^2 e^{\beta \mu N} Z_{\text{stat}}(T, V, N) = \frac{1}{\beta \Xi_{\text{stat}}} \frac{\partial}{\partial \mu} \sum_{N=1}^{\infty} N e^{\beta \mu N} Z_{\text{stat}}(T, V, N)$$

$$= \frac{1}{\beta \Xi_{\text{stat}}} \frac{\partial}{\partial \mu} (\Xi_{\text{stat}} \bar{N}) = \frac{1}{\beta} \frac{\partial N}{\partial \mu} + \frac{N}{\beta} \frac{\partial}{\partial \mu} \ln \Xi_{\text{stat}}.$$ 

Since the last term is just $\bar{N}^2$, one arrives at the result

$$\sigma_N^2 = k_B T \left( \frac{\partial N}{\partial \mu} \right)_{V,T} \quad (276)$$

In the last formula the mean $\bar{N}$ has been identified with the thermodynamical variable $N$. To bring this result to a more informative form, one imagines that $N = N(T, V, p(T, V, \mu))$ to write it as

$$\sigma_N^2 = k_B T \left( \frac{\partial N}{\partial p} \right)_{V,T} \left( \frac{\partial p}{\partial \mu} \right)_{V,T}.$$ 

Since only intensive variables are involved in the last derivative, $V$ has been replaced by $v$ - the volume per one particle. One can now use the Gibbs-Duhem differential relation (154)

$$d\mu = -s \, dT + v \, dp,$$

which implies that at constant $T$ the derivative of $p$ with respect to $\mu$ is $v$. Furthermore, since the first derivative is taken at constant volume, one can differentiate $1/v \equiv N/V$, instead of differentiating $N$:

$$\sigma_N^2 = k_B T V \left( \frac{\partial}{\partial v} \frac{1}{v} \right)_{T} \frac{1}{v} = -k_B T \frac{V}{v^2} \left( \frac{\partial v}{\partial p} \right)_{T} = k_B T \frac{v}{v^2} k_T V.$$ 

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 $1/\sqrt{N}$. 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 not be equivalent as far as the computational difficulties are concerned - this is why 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 potential of this quantity and by controlling this potential one controls the mean value of the quantity of interest in the system.

The two ensembles 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) 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.

*Quantum Grand Canonical Ensemble*

As has been explained, it is the quantum case where 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. One then imagines that the system is very weakly 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 $N$ are negligibly small (typically suppressed by the factor $1/N^{1/2}$) 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 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) **closed isolated** supersystems each consisting of a copy of the system and of the reservoir which are both **isolated and closed** i.e. not exchanging between one another particles or energy. But the distribution of the supersystem’s total energy $E_{\text{tot}}$ and of its total number $N_{\text{tot}}$ of particles between the system and the reservoir follows the principle of a priori equal weights. In this way it is possible for each particular division of $N_{\text{tot}}$ into $N_r$ and $N$ to introduce separate Hilbert spaces $\mathcal{H}_{\text{sys}}^{(N)}$ and $\mathcal{H}_{\text{res}}^{(N)}$ spanned respectively by vectors$^{120}$ $|l_1, \ldots, l_N\rangle$ and $|L_1, \ldots, L_{N_r}\rangle$ defined in (266). 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 $l_1, \ldots, l_N$ and $L_1, \ldots, L_{N_r}$. These

---

$^{120}$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 $l_1, \ldots, l_N$ (in the set $L_1, \ldots, L_{N_r}$) assume the same values.
vectors can be also written in the occupation number representation discussed in Lecture XII. In each of the Hilbert spaces $\mathcal{H}_N^{\text{sys}}$ of the system (Hilbert spaces $\mathcal{H}_{(N)_{_{\text{res}}}}$ of the reservoir) acts the corresponding Hamiltonian $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{sys}}$ of the system consisting of $N$ particles (the corresponding Hamiltonian $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{res}}$ of the reservoir consisting of $N_i = N_{\text{tot}} - N$ particles). We do not assume here that these are Hamiltonians of noninteracting systems. Still, in each of the separate Hilbert spaces $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{sys}}$ of the system ($\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{res}}$ of the reservoir) we can, as is always possible when quantum systems are not of infinite spatial extent, introduce a basis $|n_{(N)}\rangle (|n_{(N)}^{\text{res}}\rangle)$ formed by the eigenvectors of the Hamiltonian $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{sys}}$ (of the Hamiltonian $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{res}}$). If these Hamiltonians are not trivial, the eigenvectors $|n_{(N)}\rangle (|n_{(N)}^{\text{res}}\rangle)$ are built as, perhaps very complicated, linear combinations of the basis vectors $|l_1, \ldots, l_N\rangle (|L_1, \ldots, L_{N_i}\rangle)$. Only if the elements (particles) of the system (reservoir) are mutually noninteracting are the Hamiltonian eigenvectors $|n_{(N)}\rangle (|n_{(N)}^{\text{res}}\rangle)$ simply identical with the vectors $|l_1, \ldots, l_N\rangle (|L_1, \ldots, L_{N_i}\rangle)$. To each element (consisting of a copy of the system and of a copy of the reservoir) of the ensemble there corresponds therefore the Hilbert space $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{sys}} \otimes \mathcal{H}_{(N)_{_{\text{res}}}}^{\text{res}}$ which can be viewed as spaned either by the vectors

$$|l_1, \ldots, l_N\rangle \otimes |L_1, \ldots, L_{N_i}\rangle,$$

or, equivalently, by the vectors $|n_{(N)}\rangle \otimes |n_{(N)}^{\text{res}}\rangle$. It is this second basis in terms of which the assignment 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 $N_{\text{tot}} + 1$ statistical operators $\hat{\rho}_{\text{Micro}}^{(N,N_i)}$ each acting in one of the Hilbert spaces $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{sys}} \otimes \mathcal{H}_{(N)_{_{\text{res}}}}^{\text{res}}$ of the form

$$\hat{\rho}_{\text{Micro}}^{(N,N_i)} = \text{Const.} \sum_{n_{(N)}, n_{(N)}^{\text{res}}} \langle n_{(N)}\rangle \otimes \langle n_{(N)}^{\text{res}}\rangle |n_{(N)}\rangle \langle n_{(N)}| |n_{(N)}^{\text{res}}\rangle \langle n_{(N)}^{\text{res}}|$.$$

$$E_{\text{tot}} \leq E_{n_{(N)}}^{\text{sys}} + E_{n_{(N)}^{\text{res}}} \leq E_{\text{tot}} + \Delta E$$

in which

$$\text{Const.}^{-1} = \sum_{N=0}^{N_{\text{tot}}} \sum_{n_{(N)}, n_{(N)}^{\text{res}}} \langle n_{(N)}\rangle \langle n_{(N)}| |n_{(N)}\rangle \langle n_{(N)}| |n_{(N)}^{\text{res}}\rangle \langle n_{(N)}^{\text{res}}| \leq E_{\text{tot}} + \Delta E$$

The set of the statistical operators $\hat{\rho}^{(N)}$ of the Grand Canonical Ensemble pertaining to the system alone is now obtained by taking traces of the operators $\hat{\rho}_{\text{Micro}}^{(N,N_i)}$ over the Hilbert spaces $\mathcal{H}_{(N)_{_{\text{res}}}}^{\text{res}}$. Each such a trace gives

$$\rho^{(N)} \propto \sum_{n_{(N)}} |n_{(N)}\rangle \langle n_{(N)}| \exp \left\{ \frac{1}{k_B} S_{\text{res}}^{\text{res}} (E_{\text{tot}} - E_{n_{(N)}}, N_{\text{tot}} - N) \right\}.$$

Expanding now as usually the argument of the exponent

$$S_{\text{res}}^{\text{res}} (E_{\text{tot}} - E_{n_{(N)}}, N_{\text{tot}} - N) = S_{\text{res}}^{\text{res}} (E_{\text{tot}}, N_{\text{tot}}) - \frac{1}{T} E_{n_{(N)}} + \frac{\mu}{T} N + \ldots,$$

174
and discarding higher order terms of the expansion on account on the fact that they vanish in the limit of infinitely big reservoir \((N_{\text{tot}} \to \infty, E_{\text{tot}} \to \infty, E_{\text{tot}}/N_{\text{tot}} \text{ fixed})\), or modify the distribution insignificantly only if the reservoir is vary large but not really infinite, one obtains

\[
\hat{\rho}^{(N)} \propto e^{\mu N/k_B T} \sum_{n(N)} e^{-E_n(N)/k_B T} |n(N)\rangle \langle n(N)|. \tag{277}
\]

The constant of proportionality, the Grand Canonical statistical sum \(\Xi_{\text{stat}}\), is given by

\[
\Xi_{\text{stat}} = \sum_{N=0}^{\infty} e^{\mu N/k_B T} \sum_{n(N)} e^{-E_n(N)/k_B T} = \sum_{N=0}^{\infty} e^{\mu N/k_B T} \text{Tr}_{\mathcal{H}_{\text{sys}}^{(N)}}(e^{-\hat{H}_{\text{sys}}^{(N)}/k_B T}). \tag{278}
\]

As in the classical case the sum \(\Xi_{\text{stat}}(T, V, \mu)\), being related to the thermodynamic potential (we assume the system is simple) \(\Omega(T, V, \mu)\) by the formula (274), contains complete thermodynamical information about the considered system. All other characteristics of the system, like means of observables, \(O\) fluctuations around the means can be computed using the statistical operator (277).

It is practical to advance the mathematical formalism a little bit by introducing the “big” Hilbert space constructed as a direct sum

\[
\mathcal{H} = \bigoplus_{N=0}^{\infty} \mathcal{H}_{\text{sys}}^{(N)},
\]

of \(N\)-particle Hilbert spaces of the system. The Hilbert space \(\mathcal{H}_{\text{sys}}^{(0)}\) corresponding to zero particles is here an artificially constructed one-dimensional vector space spanned by a single vector \(|\text{Void}\rangle\). General vectors \(|\Psi\rangle\) of \(\mathcal{H}\) have, of course components in all subspaces \(\mathcal{H}_{\text{sys}}^{(N)}\)

\[
|\Psi\rangle = c_0|\text{Void}\rangle + \sum_{N=1}^{\infty} c_N |\Psi^{(N)}\rangle,
\]

where each vector \(|\Psi^{(N)}\rangle\) is a linear combination of the basis vectors (266). 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 \(\mathcal{H}_{\text{sys}}^{(N)}\)) 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\(^{121}\)

\[
a_l^\dagger |l_1, \ldots, l_N\rangle = |l, l_1, \ldots, l_N\rangle,
a_l |l_1, \ldots, l_N\rangle = \sum_{j=1}^{N} \zeta^{j-1} \delta_{l,l_j} |l_1, \ldots, (\text{no } l_k, \ldots, l_N)\rangle, \tag{280}
\]

\(^{121}\)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 (206).
where \( \zeta = 1 \) in the case of bosons and \(-1\) is the particles 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

\[
[a_\nu, a_l^\dagger] = \delta_{\nu l}, \quad [a_\nu, a_l] = [a_\nu^\dagger, a_l^\dagger] = 0, \tag{281}
\]

if the particles are bosons and the anticommutation rules

\[
\{a_\nu, a_l^\dagger\} = \delta_{\nu l}, \quad \{a_\nu, a_l\} = \{a_\nu^\dagger, a_l^\dagger\} = 0, \tag{282}
\]

where \( \{\cdot, \cdot\} \) is the anticommutator, if the particles are fermions. In the occupation number representation \( |n_1, n_2, \ldots\rangle \) of the basis vectors (266) of the action of the bosonic operators takes the familiar (at least to some) form

\[
\begin{align*}
a_l^\dagger |n_1, \ldots, n_l, \ldots\rangle &= \sqrt{n_l + 1} |n_1, \ldots, n_l + 1, \ldots\rangle, \\
a_l |n_1, \ldots, n_l, \ldots\rangle &= \sqrt{n_l} |n_1, \ldots, n_l - 1, \ldots\rangle. \tag{283}
\end{align*}
\]

The action of the fermionic operators in this representation is less familiar

\[
\begin{align*}
a_l^\dagger |n_1, \ldots, n_l, \ldots\rangle &= \begin{cases} 0 & \text{if } n_l = 1 \\ \eta |n_1, \ldots, 1, \ldots\rangle & \text{if } n_l = 0 \end{cases} \\
a_l |n_1, \ldots, n_l, \ldots\rangle &= \begin{cases} \eta |n_1, \ldots, 0, \ldots\rangle & \text{if } n_l = 1 \\ 0 & \text{if } n_l = 0 \end{cases}. \tag{284}
\end{align*}
\]

\( \eta \) is here some phase factor which we will not need (see my Lectures on Quantum Field Theory).

Usefulness of the creation and annihilation operators stems from the fact that any operator acting in individual \( N \)-particle Hilbert spaces \( \mathcal{H}_{\text{sys}}^{(N)} \) can be expressed through them and at the same time promoted to act in the whole big space \( \mathcal{H} \).