PHYSICS
262

STATISTICAL
MECHANICS

PHYSICS 262

THERMODYNAMICS AND STATISTICAL MECHANICS

INSTRUCTOR: GLAUBER — ROOM J256: TTS12

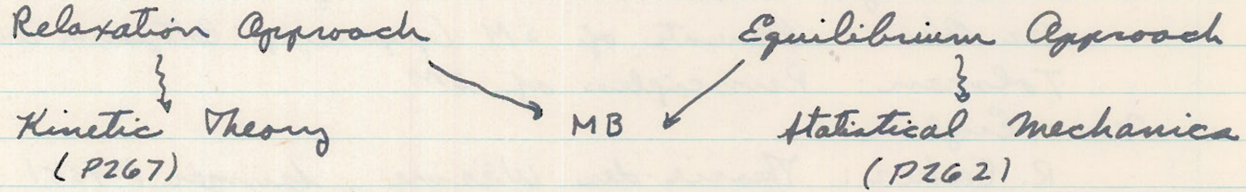
LECTURE 1: 9-26-61

Historical Development:

Clausius } 1859
Maxwell }

Maxwell used a heuristic differential equation to derive the MB (Maxwell-Boltzmann) distribution. Not until Boltzmann (1860-80) did a serious attempt at derivation take place. He originally used the time dependent approach which described the relaxation of a system into MB form from some initial velocity distribution with the passage of time. However, criticism of this method at the time led Boltzmann to find another approach which later became the basis of statistical mechanics and which was called the equilibrium approach. This approach assumes no macroscopic motion, only microscopic. It used results of probability theory. If we single out a single particle at a time repeatedly, it is found that the most probable velocity distribution is MB. He used the concept of phase space in his derivation.

On the other hand, Boltzmann's two developments unfortunately gave rise to two schools of thought:



However, even SM (statistical mechanics) uses time dependence approach when it asks for the probability of $\vec{v}$ at time t knowing $\vec{v}_0$ initially, $P(\vec{v}_0, \vec{v}\, t)$.

In the equilibrium approach, very little knowledge is needed of the details of the system. That is, the essentials of equilibrium theory need not consider whether a system is a solid, liquid, or gas. SM is a complete theory in that one can formulate all problems of a physically statistical nature in terms of it. Even quantum mechanics falls into place with SM very well. At present, SM can be formulated by recourse to several theorems about systems with many degrees of freedom.

We will not use Boltzmann's approach to the MB distribution as it is contained in many texts and does not evaluate the errors inherent in its derivation.

Requisite knowledge: Classical Mechanics
Quantum mechanics
Thermodynamics (self-contained in course)
Mathematics: contour integration, δ functions

Textbooks and References:

Fowler; Statistical Mechanics (classic)
Gibbs: Elementary Principles of SM (1902)
Hill: SM (1956)
Khinchin: Mathematical Foundations of SM (Russian, Dover)
Landau & Lifshitz: Statistical Physics
Kittel: Elementary SM
Mayer & Mayer: SM
Schroedinger: Statistical Thermodynamics (worthwhile)
ter Haar: Elements of SM (papers of Boltzmann and Ehrenfest)
Tolman: Principles of SM

Non-English:

R. Becker: Theorie der Wärme, Springer (1956)
Chintschin: Mathematische Grundlagen der Quantenstatistik
Munster: Statistische Thermodynamik
Notes of K. Huang at MIT course last year.

Classical Mechanical Foundation:

Imagine a dynamical system with F degrees of freedom described by the coordinates:

$$q_k ; \dot{q}_k \quad ; \quad k = 1, \ldots, F$$

The system can be described completely in these coordinates. Recall the Lagrangian $\mathcal{L}(q_k, \dot{q}_k, t)$ and the subsequent equation of motion:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k}\right) = \frac{\partial \mathcal{L}}{\partial q_k}$$

We can formulate an alternative description of the motion by the use of the Hamiltonian. Introduce the generalized momentum:

$$p_k \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_k}$$

Then the Hamiltonian is defined by:

$$H(q_k, p_k, t) = \sum_{k=1}^{F} \left[p_k \dot{q}_k - \mathcal{L}(q_k, \dot{q}_k, t) \right]$$

Now vary H with respect to all of the variables, p_k, q_k, $\dot{q}_k$, t:

$$dH = \sum_k \left(p_k \, d\dot{q}_k + \dot{q}_k \, dp_k \right) - \sum_k \left(\frac{\partial \mathcal{L}}{\partial q_k} dq_k + \frac{\partial \mathcal{L}}{\partial \dot{q}_k} d\dot{q}_k \right) + \frac{\partial \mathcal{L}}{\partial t} dt$$

Using the definition of the generalized momenta, we see that two terms cancel and we can write:

$$\frac{\partial H}{\partial p_k} = \dot{q}_k \quad ; \quad -\frac{\partial H}{\partial q_k} = \dot{p}_k \quad ; \quad -\frac{\partial H}{\partial t} = \frac{\partial \mathcal{L}}{\partial t}$$

These are Hamilton's equations of motion.

Now consider some arbitrary function of the generalized coordinates:

$$\mathcal{G}(q_k, p_k, t)$$

suppose we have already solved for q_k, p_k as functions of time. We then ask what is the time derivative of G? From partial differentiation:

$$\frac{d}{dt} G = \sum_k \left[\frac{\partial G}{\partial q_k}\frac{dq_k}{dt} + \frac{\partial G}{\partial p_k}\frac{dp_k}{dt} \right] + \frac{\partial G}{\partial t}$$

$$= \underbrace{\sum_k \left[\frac{\partial G}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial G}{\partial p_k}\frac{\partial H}{\partial q_k} \right]}_{\text{Poisson Brackets}} + \frac{\partial G}{\partial t}$$

$$= \left[G, H\right] + \frac{\partial G}{\partial t}$$

Note the striking similarity to Heisenberg's equation of motion in quantum mechanics.

Consider p_k, q_k as a set of $2F$ variables, indexed by $j = 1, \dots, 2F$. Define $2F$ variables x_j as follows:

$$x_k = q_k \quad ; \quad k = 1, \cdots, F$$
$$x_{F+k} = p_k \quad ; \quad k = 1, \cdots, F$$

The $2F$ points is said to form a <u>phase space</u> for the system. Now define a velocity function in this space:

$$v_k = \frac{\partial H}{\partial p_k} \quad ; \quad k = 1, \cdots, F$$

$$v_{F+k} = -\frac{\partial H}{\partial q_k} \quad ; \quad k = 1, \cdots, F$$

Hence the equation of motion in the phase space is:

$$\dot{x}_j = v_j(x_1, \dots x_{2F}, t) \quad ; \quad j = 1, \cdots, 2F$$

We take up next time trajectories in phase space.

LECTURE 2: 9-28-61

Recapitulation: $\dot{q}_k = \frac{\partial H}{\partial p_k}$; $x_j = q_j$, $1 \le j \le F$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} \; ; \quad x_j = p_{j-F}, \; F+1 \le j \le 2F$$

These relations allow us to cast the equations of motion into new form:

$$\dot{x}_j = v_j(x_1 \cdots x_{2F}, t)$$

In the phase space, this is a coupled differential equation for the velocity of the phase points. An unique solution is implied which leads to the construction of trajectories in the phase space. Hence:

$$x_j(t) = X_j(x_1^{(0)} \cdots x_{2F}^{(0)}, t)$$

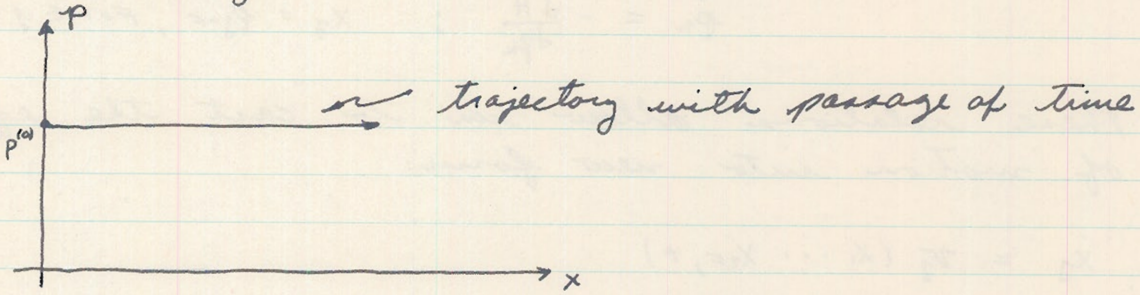
where $^{(0)}$ denotes an initial condition. Usually there are constraints placed on the phase space by considerations of maximum energy, volume, etc. Every accessible point in PS (phase space) is suitable for an initial condition from which we can start. Each point has a single trajectory which passes through it, and transforms into another unique point with the passage of time.

Can two trajectories intersect? This is really a question of uniqueness. If intersection be possible, a point has a chance to go in two different directions at the point of intersection. This is not possible because of the uniqueness of the integration of the equation of motion.

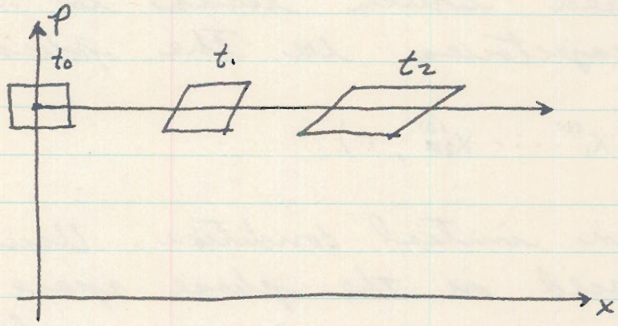
In general, we deal with conservative systems which do not depend explicitly on the time.

Suppose we a volume in PS as an initial condition. Solution of the equations of motion give a transformation into a new shape. Often we must talk about a volume in phase space because we do not know the initial condition of a point exactly.

Consider a single particle with initial conditions $x^{(0)}, p^{(0)}$. Take particle free and not subject to external forces. Classically, if we know $x^{(0)}, p^{(0)}$ we know future motion exactly.



What if we are uncertain about the values of $x^{(0)}, p^{(0)}$ except that they lie in a certain volume?



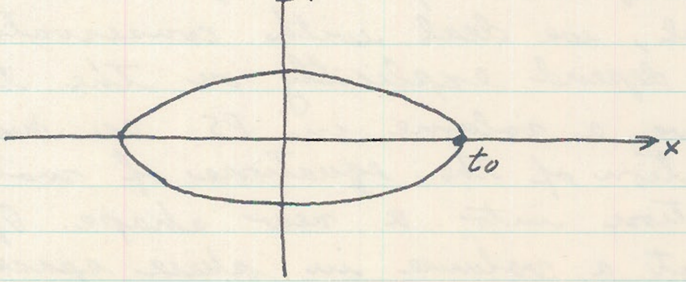
The volume changes shape with time because points with greater $p^{(0)}$ go further in a given time interval.

The uncertainty in position and momentum introduces a correlation between the two.

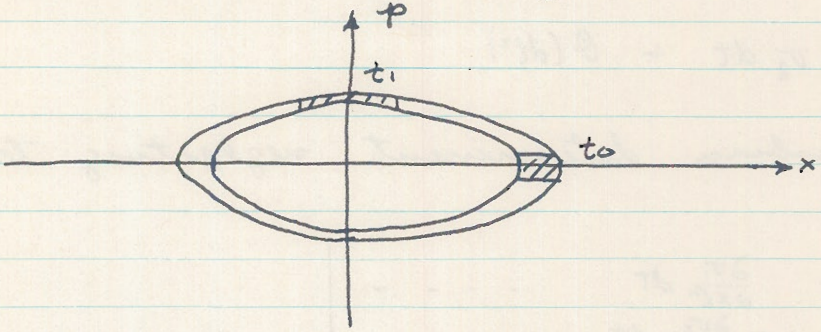
Consider another example; The Harmonic Oscillator:

$$H = \frac{p^2}{2m} + \frac{1}{2} k x^2 = E$$

Start at rest and stretch to $x^{(0)}$, $p^{(0)} = 0$ and release. The trajectory in PS or surface of constant energy will be an ellipse.



Now suppose an uncertainty in the initial conditions:



The shape changes, among other things, to it new shape according to k and m.

Liouville Theorem:

This theorem states that the volume of a set of points in PS remains unchanged by changes induced in the dynamical motion. For example, the area concerned in the two previous problems remains unchanged with time.

Proof:

Consider the initial volume at t_0 described by the set of points ω_0. We say that it transforms in time from the set ω_0 to the set ω_t. Consider the volume of the set ω_t:

$$V(\omega_t) = \int_{\omega_t} dx_1 \cdots dx_{2F}$$

We now transform to an integral taken over ω_0 by using the equations of motion.

$$V(\omega_t) = \int_{\omega_0} \underbrace{\frac{\partial(x_1 \cdots x_{2F})}{\partial(x_1^{(0)} \cdots x_{2F}^{(0)})}}_{\text{Jacobian}} dx_1^{(0)} \cdots dx_{2F}^{(0)}$$

The problem then resolves to finding the Jacobian (J). We now consider a time t very close to t_0 and expand the solution of the equation of motion for x_j:

$$x_j(t) = x_j^{(0)} + \dot{x}_j(t-t_0) + \cdots = x_j^{(0)} + v_j(t-t_0) + \cdots$$

Since $t - t_0$ is very small, make it dt and write:

$$x_j(t) = x_j^{(0)} + v_j\,dt + \mathcal{O}(dt^2)$$

Set up the Jacobian determinant neglecting terms of $\mathcal{O}(dt^2)$:

$$\begin{vmatrix} 1 + \frac{\partial v_1}{\partial x_1^{(0)}}\,dt & \frac{\partial v_1}{\partial x_2^{(0)}}\,dt & - - - - \\ \frac{\partial v_2}{\partial x_1^{(0)}}\,dt & 1 + \frac{\partial v_2}{\partial x_2^{(0)}}\,dt & \\ \vdots & & \ddots \end{vmatrix}$$

How do we evaluate this? We can form products from the terms in each row and column $N!$ ways. Hence, one of the resulting terms will be the product of the diagonal terms:

$$\big|\;\big| = \prod_{j=1}^{2F}\left(1 + \frac{\partial v_j}{\partial x_j^{(0)}}\,dt\right) + \cdots$$

Since we have used up all the diagonal terms, any products involving an off-diagonal term must also involve at least one other and hence be of $\mathcal{O}(dt^2)$. Also, in the diagonal product, only crossproducts with one will find themselves of first order, hence:

$$\big|\;\big| = 1 + \sum_{j=1}^{2F} \frac{\partial v_j}{\partial x_j^{(0)}}\,dt + \mathcal{O}(dt^2)$$

Note that $\sum_{j=1}^{2F} \frac{\partial v_j}{\partial x_j^{(0)}}$ has the form of a divergence in PS.

What is the value of this divergence? We are talking of the divergence of some velocity field in PS. Use Hamilton's relations:

$$\sum_{j=1}^{2F} \frac{\partial v_j}{\partial x_j^{(0)}} = \sum_{k=1}^{F}\left\{\frac{\partial}{\partial q_k^{(0)}}\frac{\partial H}{\partial p_k^{(0)}} - \frac{\partial}{\partial p_k^{(0)}}\frac{\partial H}{\partial q_k^{(0)}}\right\} = 0$$

because of the nature of mixed partials.

Thus the divergence of the velocity field in PS is zero. We now can write:

$$\Big|\ \Big| = \frac{\partial(x_1 \cdots x_{2F})}{\partial(x_1^{(0)} \cdots x_{2F}^{(0)})} = 1 + \Theta(dt^2)$$

Thus we find:

$$\frac{d}{dt}\big|\,\big| = 0$$

The dt^2 terms do not contribute anything as the first derivative of a second order differential vanishes. Thus the above proof is rigourous and does not hold to only first order in time.

We now offer an alternative approach. We gave one method of showing that the Jacobian is unity in the first order of time and does not change with time for all order of time. If we solve $\frac{d}{dt}|\,| = 0$, we get $|\,|$ = to its initial value at t_0 which is unity. Hence $|\,| \equiv 1$ and the initial volume in phase space does not change with time. Consider an alternative approach. Define for finite time:

$$J(t) \equiv \frac{\partial(x_1 \cdots x_{2F})}{\partial(x_1^{(0)} \cdots x_{2F}^{(0)})}$$

and find $\frac{d}{dt} J(t)$.

Since $|J(t)|$ is a sum of products, we can write a determinant as a sum of determinants in the following way:

$$\frac{d}{dt} J(t) = \sum_j \frac{\partial(x_1 \cdots x_{j-1}, \dot{x}_j, x_{j+1} \cdots x_{2F})}{\partial(x_1^{(0)} \cdots x_{2F}^{(0)})}$$

$$= \sum_j \frac{\partial(x_1 \cdots x_{j-1}, v_j, x_{j+1} \cdots x_{2F})}{\partial(x_1^{(0)} \cdots x_{2F}^{(0)})}$$

Now we can form the identity:

$$\frac{\partial v_j}{\partial x_n^{(0)}} = \sum_{l=1}^{2F} \frac{\partial v_j}{\partial x_l} \frac{\partial x_l}{\partial x_n^{(0)}}$$

Then:

$$\frac{d}{dt} J(t) = \sum_{j=1}^{2F} \sum_{l=1}^{2F} \frac{\partial v_j}{\partial x_l} \frac{\partial (x_1 \cdots x_{j-1}, x_l, x_{j+1} \cdots x_{2F})}{\partial (x_1^{(0)} \cdots x_{2F}^{(0)})}$$

We see that this new determinant vanishes most of the time unless $l = j$ because a row is bound to be repeated otherwise. Thus:

$$\frac{\partial (x_1 \cdots x_{j-1}, x_l, x_{j+1} \cdots x_{2F})}{\partial (x_1^{(0)} \cdots x_{2F}^{(0)})} = \delta_{jl} \, J(t)$$

Therefore:

$$\frac{d}{dt} J(t) = \sum_{j=1}^{2F} \frac{\partial v_j}{\partial x_j} J(t) = 0$$

because we have shown that the divergence of the velocity in PS vanishes. This then is the equation of motion for $J(t)$ as one moves along a dynamical trajectory and verifies our previous result. Finally:

$$V(\omega_t) = \int_{\omega_0} J \, dx_1^{(0)} \cdots dx_{2F}^{(0)} = \int_{\omega_0} dx_1^{(0)} \cdots dx_{2F}^{(0)} = V(\omega_0) \text{, QED}$$

A given volume in phase space remains unchanged with time. This is a monumentally general statement regardless of interactions, external fields, etc. It is also a triumph of the Hamiltonian formulation. If we had use v instead of p, this would not always be true. Consider the magnetic field case where v is not simply p/m.

$$v = \frac{1}{m} (p - \frac{e}{c} A)$$

Hence it is the use of the canonical momentum that makes the theorem so general.

note: in textbooks, the above is not called the Liouville theorem, there they talk about motion of some density function. The Liouville theorem here is the way Liouville meant it and is not statistical and applies to the entire system. In summary, we can state mathematically that the measure of a system is conserved or preserved by a dynamical transformation.

Consequences of the Liouville Theorem:

If a system is confined in a volume in PS, in that volume history repeats itself. Whatever point we begin from, the subsequent motion is bound to return us to an arbitrarily close neighborhood of the initial point after an arbitrarily long time. Say we start with the particles of a gas arranged with an initial position and momenta, sooner or later, the velocity and position of the particles will return as close as we like to the initial point. This is called <u>Poincaré's Recurrence Theorem</u>.

Consider a volume in PS and then how it transforms in time. Then consider a trajectory constructed on each point in the initial volume. All trajectories except those of measure zero will eventually return to the initial volume. Physically, the term measure zero means all will return with probability one after a long enough time interval.

LECTURE 3: 10-3-61

Liouville Theorem: Poincaré Recurrence Theorem:

We speak only of a system confined to a limited region of phase space. This can be done by limiting the total energy of the system or confining its volume. Consider a point moving in P.S. The statement of the theorem is: for almost all initial points, the trajectory will bring you back as close to the initial point as you want if you wait long enough. (Mathematically all most all, physically all)

Proof: (due to Carathéodory)

Consider a volume of PS, w_0. This volume is carried into successive volumes with time. We need not consider time continuous (remarkable). We consider the system at intervals $0, t, 2t, \ldots$ Consider point $P \in w_0$ ($\in \rightarrow$ belonging to). We then consider the dynamical transformation $\mathcal{T}$ upon P: $\mathcal{T}P$. $\mathcal{T}P$ takes the system to time t, $\mathcal{T}^2P$ to $2t$, etc.

Nearly every point P in w_0 has the property that one of the transformations $\mathcal{T}$ ($\sim\mathcal{T}^{10^{10}}$ most likely) operating on it returns it to its initial state. We can choose w_0 arbitrarily small so that points can be brought back as close as possible

What about points that don't return? Call S the subset of points P such that there is no return to w_0 during dynamical sequence. We want to show that the measure of this subset is zero.

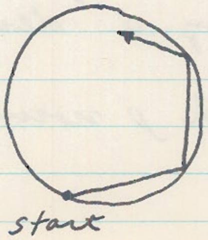
Consider a dynamical transformation applied to all <u>points in S</u> and further consider more, that is: $S, \mathcal{T}S, \mathcal{T}^2S$, etc. We thus generate a sequence of sets which we stipulate have no overlap. We now invoke the Liouville Theorem in stating that all these sets have identical volume or measure. Somehow these must be fitted into a finite volume in PS, finite because of restraints (energy, volume).

To show that there is no overlap, suppose that that sets $\mathcal{T}^j S$ and $\mathcal{T}^k S$, $k > j$, do have points in common. Because we can run back in time, the dynamical transformations have a unique inverse. Apply $(\mathcal{T}^{-1})^j$ on $\mathcal{T}^j S$ and $\mathcal{T}^k S$. Then S and $\mathcal{T}^{k-j} S$ must have points in common but this contradicts the hypothesis that S never returns to ω_0 because we see the operation $\mathcal{T}^{k-j} S$ would return the overlap points to ω_0. Therefore there is no overlap.

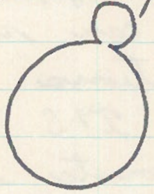
As stated before, the Liouville theorem says that all the sets have identical volume or measure and somehow must be fitted into a finite volume in PS. The only way this can be done and still keep the motion confined to a finite volume is for S to have zero measure. This means that S has a manifold of at least one below the number of dimensions of the PS because a manifold of a set has zero volume in that set (In 3-D, area has zero volume, line has zero volume).

Thus, all points that count return to the original state or we call the motion quasi-periodic because it does not necessarily return periodically. In a finite universe, this gives rise to a "reincarnation" cosmology.

What is an example of an exception? The example is somewhat artificial: Consider a particle bouncing around in a hollow sphere. The particle keeps making unclosed polygons but after a sufficient time, it would return arbitrarily close to its initial position and momentum. Of course if the polygon is closed, it will return periodically.



now attach another sphere with a common hole between them of some finite size.



The particle will sometime bounce into the other sphere but will also return so it will eventually get back to its initial state. If we close down to a mathematical point, only a few initial positions will cause the particle to pass thru and it will not return until after an infinite time or it does not return with measure zero.

Further significance of the recurrence theorem:

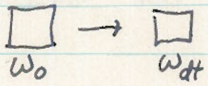
One of the fundamental processes of thermodynamics, or fundamental notions rather, is that of irreversible processes. An exploding balloon is an example. However, we have shown the laws of mechanics to be reversible since they arise from a second order differential equation. Boltzmann derived his relation from a first order differential equation hence the origin of irreversibility. The catch is that one must wait nearly infinitely long for recurrence to occur in mechanics. However, irreversible thermodynamics is a good approximation over a short time.

We now consider the concept of <u>probability density</u> in PS. Recall the position of a point in phase space: $x_1 \cdots x_{2F}$. However, we rarely know the precise state of a system, so we represent the probable state by a density function:

$$\rho(x_1 \cdots x_{2F}) \quad ; \text{ with } \rho \text{ normalized: } \int \rho(x_1 \cdots x_{2F}) \prod_j dx_j = 1$$

ρ represents what we know about the system and also what we don't know. However, to determine ρ we have to know it as a function of about 10^{23} variables. Thus it is seldom used except in an epistemological sense. However, let us consider how we would find ρ. We could keep making repeated measurements on a prepared system or repeat the same measurement on a number of freshly prepared systems. Eventually we would obtain a collection of points in PS which would determine ρ. This is the Gibbs' ensemble used by Gibbs because of a distatse for probability. However it describes the same things as ρ does so that we will not use the concept of Gibbsian ensembles but in the spirit of the course we will use ρ.

What about the time dependence of ρ? : $\rho(x_1 \cdots x_{2F}, t)$ Consider the infinitesimal cloud of points ω_0 going into ω_{dt} in time dt.



Consider $\rho_0 V(\omega_0)$ as number of points in $V(\omega_0)$. By conservation of mass principle:

$$\rho_0 V(\omega_0) = \rho_{dt} V(\omega_{dt})$$

The Liouville Theorem demands $V(\omega_0) = V(\omega_{dt})$, hence $\rho_0 = \rho_{dt}$ or as we follow along a trajectory:

$$\frac{d\rho}{dt} = 0$$

This is sometimes called the Liouville Theorem in the texts. We will call it the <u>Liouville Equation</u>.

Consider for a moment what $\frac{d}{dt}$ really is. In hydrodynamics we have what is called the <u>substantial derivative</u>:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \vec{v}\cdot\nabla = \frac{\partial}{\partial t} + \sum_j v_j \frac{\partial}{\partial x_j} \quad \text{, in PS notation}$$

Then for ρ:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_j v_j \frac{\partial}{\partial x_j}\rho = 0$$

or $$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{\lambda=1}^{F}\left\{\frac{\partial \rho}{\partial q_\lambda}\frac{\partial H}{\partial p_\lambda} - \frac{\partial \rho}{\partial p_\lambda}\frac{\partial H}{\partial q_\lambda}\right\} = 0$$

or we can state the <u>Liouville Equation</u> in the Poisson Bracket notation:

$$\frac{\partial \rho}{\partial t} + [\rho, H] = 0$$

Another approach is from the point of view of the Continuity Theorem of hydrodynamics: We think of a fluid with density ρ at velocity $\vec{v}$. The continuity equation gives:

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\vec{v}) = 0$$

or $$\frac{\partial \rho}{\partial t} + \vec{v}\cdot\nabla\rho + \rho\,\nabla\cdot\vec{v} = 0$$

We see that $\nabla\cdot\vec{v} = 0$ which in hydrodynamics characterizes incompressible flow. This does not mean the density is uniform, but means the density around each point remains constant with time. A consequence of the Continuity Equation or the Liouville Equation is that normality is preserved.

If: $$\int \rho(x_1 \cdots x_{2F}, 0)\prod_j dx_j = 1$$

Then: $$\int \rho(x_1 \cdots x_{2F}, t)\prod_j dx_j = 1$$

Proof:

$$\frac{\partial}{\partial t}\int \rho \prod_j dx_j = -\int \nabla\cdot(\rho\vec{v}) \prod_j dx_j$$

$$= -\int_{\substack{\text{surface}\\ \text{in PS}}} \rho\vec{v}\cdot d\vec{S} = 0$$

The surface integral vanishes if the fluid is at all confined because $\rho = 0$ at an infinite distance and can be made to vanish on an infinite surface. This is always true as far as we are concerned.

Facts about the Liouville Equation:

It contains the same information as the equations of motion, that is, if we specify ρ, we can have the same information that is contained in the equations of motion. If we know the initial conditions at $t=0$, $x_j^{(0)}$, then at $t=0$, the description of ρ is $\rho(x_1 \cdots x_{2F}, 0)$ written as:

$$\rho(x_1 \cdots x_{2F}, 0) = \prod_{j=1}^{2F} \delta(x_j - x_j^{(0)})$$

Recall: $\delta(x) = 0,\ x \neq 0,\ \delta(0) = \infty,\ \int \delta(x)\,dx = 1$

What about subsequent times?

$$x_j(t) = X_j(x_1^{(0)} \cdots x_{2F}^{(0)}, t)$$

Then for ρ at time t we stipulate the form:

$$\rho(x_1 \cdots x_{2F}, t) = \prod_{j=1}^{2F} \delta\left[x_j - X_j(x_1^{(0)} \cdots x_{2F}^{(0)}, t)\right]$$

Proof: $\frac{\partial \rho}{\partial t} = \sum_k \prod_{j\neq k} \delta(x_j - X_j)(-v_k)\,\delta'(x_k - X_k)$

$$\sum_k v_k \frac{\partial}{\partial x_k}\rho = \sum_k \prod_{j\neq k} \delta(x_j - X_j)\, v_k\, \delta'(x_k - X_k)$$

Thus the assumed solution for ρ satisfies the Liouville equation and shows the density around a point in PS stays concentrated with time.

LECTURE 4 : 10-5-61

Recapitulation: definition of density function: $\rho(x_1 \cdots x_{2F}, t)$ We can get equation of motion of ρ from Liouville equation or continuity equation if $\nabla \cdot \vec{v} = 0$.

Suppose we are interested in the typical values of some function of the PS. We call this:

$$\mathcal{G}(q_1 \cdots q_F, p_1 \cdots p_F, t)$$

One way to get typical values is to take the average with respect to ρ. Define, then, as the average value of $\mathcal{G}$:

$$\bar{\mathcal{G}}(t) = \int \mathcal{G}(q_1 \cdots p_F, t)\, \rho(q_1 \cdots p_F, t) \prod_{j=1}^{F} dq_j\, dp_j$$

$\bar{\mathcal{G}}$ will be the experimentally observed quantity. If we choose a set of identically prepared systems, ensemble, and make a measurement of $\mathcal{G}$ at time t of each system and average, we get the ensemble average. This is the average we discuss in this course. This is the only kind we calculate or talk about.

We would still have quite a problem working with ρ in all its glory. Even if we had ρ as a function of all the PS variables, it would be almost useless. In practice almost all information in ρ goes to waste because the actual function $\mathcal{G}$ we work with only involves a very few coordinates. An example is the kinetic energy of a system:

$$T = \sum_j \frac{p_j^2}{2m_j}$$

$$\bar{T} = \int \sum_j \frac{p_j^2}{2m_j}\, \rho(q_1 \cdots p_F, t) \prod_{k=1}^{F} dq_k\, dp_k$$

For any given particle of momentum p_j, we have just the job of integrating out the rest of the variables.

We thus define the single particle density function:

$$\rho_j^{(1)}(q_j, p_j) = \int \rho(q_1 \cdots p_F) \prod_{k \neq j} dq_k\, dp_k$$

so most of the information is swept out with the integration. Then:

$$\overline{T} = \sum_j \int \frac{p_j^2}{2m_j} \rho_j^{(1)}(q_j\, p_j)\, dq_j\, dp_j$$

So this is one example of the reduction of the problem. What about the Potential Energy? Assume not dependent on p. Assume further that we have a field of coulomb interaction between pairs of particles. Take q's as cartesian coordinates:

$$V(q_1 \cdots q_F) = \sum_{j<k} \varphi(|q_j - q_k|)$$

If the forces depended upon the position of many particles, the problem would be complicated. Gravitational and electric fields only interact between pairs at a time regardless of the position of others. Define the two-particle ρ:

$$\rho_{jk}^{(2)}(q_j q_k p_j p_k) = \int \rho(q_1 \cdots p_F) \prod_{l \neq j,k} dq_l\, dp_l \quad \text{, Therefore:}$$

$$\overline{V} = \sum_{j<k} \int \varphi(|q_j - q_k|)\, \rho_{jk}^{(2)}(q_j q_k p_j p_k)\, dq_j\, dq_k\, dp_j\, dp_k$$

We will always be able to find the thermodynamic quantities from the one or two particle ρ's.

Another simple example: number of particles in an elementary volume: Consider a subvolume. How many particles N_V in volume V. Define function:

$$\psi(q) = \begin{cases} 1 & q \text{ in } V \\ 0 & q \text{ outside } V \end{cases}$$

Ask whether each particle is in or out.

This is tantamount to finding the average value of N_ν:

$$\bar{N}_\nu = \sum_j \int \Psi(q_j)\, \rho^{(1)}(q_j p_j)\, dq_j\, dp_j$$

Assigned Problem ①:

Consider $\mathcal{G}(q_1 \cdots p_F, t)$ and $\bar{\mathcal{G}}(t)$ and $\frac{d}{dt}\bar{\mathcal{G}}(t)$. Show that:

$$\frac{d}{dt}\left(\bar{\mathcal{G}}(t)\right) = \overline{\left(\frac{d\mathcal{G}}{dt}\right)}$$

where $\frac{d\mathcal{G}}{dt}$ is a derivative which follows motion through a phase point. Consider $\rho(q_1 \cdots p_F, t)$ depending on everything:

Systems in Equilibrium:

Many particle systems have vast complexity, we cannot observe all detail. However, it is possible to find a state seen from the lab scale that apparently is not changing. This is equilibrium.

The idea of Thermodynamic isolation is quite an idealization. Equilibrium involves some kind of interaction with the outside world which brings it about. The only way to define equilibrium is to watch and see when system stops changing. Then any initial knowledge prior to equilibrium is useless. Once in equilibrium, knowledge is somewhat independent of time. What can we say about ρ? We can say it is independent of time because in equilibrium the system does not change. The assumption of ρ independent of time for equilibrium is the hypothesis that underlies all future developments on equilibrium problems. How then does it satisfy the Liouville equation? As long as ρ is constant along a trajectory, the Liouville equation is satisfied. Such quantities constant along a trajectory are called constants of the motion. Consider some constants of the motion:

$$\alpha(q_1 \cdots q_F) = \text{constant}$$

$\rho(\alpha) = \rho[\alpha(q_1 \cdots q_F)]$ satisfies the Liouville equation.

The Poisson Bracket (PB) of $\mathcal{I}$ and H is:

$$[\mathcal{I}, H] = \frac{d\mathcal{I}}{dt} - \frac{\partial \mathcal{I}}{\partial t} \; ; \quad \text{but } \frac{\partial \mathcal{I}}{\partial t} = 0$$

Furthermore, since $\mathcal{I}$ is a constant of the motion, $\frac{d\mathcal{I}}{dt} = 0$ (constant along trajectory), then $[\mathcal{I}, H] = 0$

Now $\rho = \rho(\mathcal{I})$: $\quad \frac{\partial \rho}{\partial t} = -[\rho, H] = -\frac{d\rho}{d\mathcal{I}}[\mathcal{I}, H] = 0$

so that our choice of $\rho = \rho(\mathcal{I})$ gives us:

$$\frac{\partial \rho}{\partial t} = 0 \quad \text{for all fixed } q_1 \cdots p_F$$

This is not the only solution of the Liouville equation, That is, whatever we choose for $\mathcal{I}$, as there are other integrals of motion. Recall:

$$\frac{\partial H}{\partial q_k} = -\dot{p}_k \; ; \quad \frac{\partial H}{\partial p_k} = \dot{q}_k \qquad 2F \text{ equations}$$

Form:

$$\frac{\frac{\partial H}{\partial q_k}}{\frac{\partial H}{\partial p_k}} = -\frac{dp_k}{dq_k} \quad \text{which runs along a trajectory.}$$

Suppose now, we try to put the time into just one equation. Single out:

$$\frac{\partial H}{\partial p_F} = \dot{q}_F$$

$$\frac{\frac{\partial H}{\partial q_k}}{\frac{\partial H}{\partial p_F}} = -\frac{dp_k}{dq_F} \; ; \quad k = 1 \cdots F$$

$$\frac{\frac{\partial H}{\partial p_k}}{\frac{\partial H}{\partial p_F}} = \frac{dq_k}{dq_F} \; ; \quad k = 1 \cdots F-1$$

Thus we have taken out the time dependence in $2F-1$ of the equations and left it in: $\frac{\partial H}{\partial p_F} = \dot{q}_F$

This means that we can form 2F-1 integrals of the motion:

$$\left.\begin{array}{l} I_1(q_1 \cdots p_F) = C_1 \\ I_{2F-1}(q_1 \cdots p_F) = C_{2F-1} \\ I_{2F}(q_1 \cdots p_F) = C_{2F} + t \end{array}\right\} \text{2F equations of motion}$$

We see that C_{2F} is origin of time. The first 2F-1 equations determine intersecting surfaces or the trajectory. In view of this, the most general time independent ρ is:

$$\rho = \rho\left[I_1(\cdots)\, I_2(\cdots) \cdots I_{2F-1}(\cdots) \right]$$

$$\text{with} \quad \frac{\partial \rho}{\partial t} = 0 \,, \quad [\rho, H] = \sum_j \frac{\partial \rho}{\partial I_j} [I_j, H] = 0$$

This is the most general static density. It says that ρ may vary from trajectory to trajectory in PS but is constant along each. A complication might arise if the C's are multiple-valued. The C's give a unique trajectory, but I may not give a unique C. The constant of motion we will talk most about is the energy of an isolated system. To obtain equilibrium we had to have contact with some thermal body, now considered removed and the system isolated. Thus one of the C's is the energy, call it C_1. Motion then takes place in PS on a constant energy surface. Now H is the energy or:

$$H(q_1 \cdots p_F) = C_1 \quad \text{describes the constant energy surface.}$$

Hence one stationary density would be:

$$\rho\left[H(q_1 \cdots p_F)\right]$$

H <u>must</u> be independent of time for equilibrium and we see that it satisfies the Liouville equation.

now examine the continuity equation:

$$\frac{\partial\rho}{\partial t} + \vec{v}\cdot\nabla\rho = 0$$

now $\vec{v}$ is tangent to an energy surface and:

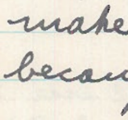
$\nabla\rho = \frac{d\rho}{dH}\nabla H$ is normal to a constant energy surface.

$\therefore\ \vec{v}\cdot\nabla\rho = 0$ as it should be.

List of some constants of the motion:

$$\left.\begin{array}{l}\text{Energy}\\ \text{Linear momentum}\\ \text{Angular momentum}\end{array}\right\}\ \begin{array}{l}\text{together 7 famous constants}\\ \text{of the motion.}\end{array}$$

They are famous because of their connection with invariance laws. When a differential equation is invariant under time, we have an integral of energy. For linear invariance, linear momentum integral; angular invariance, angular momentum integral. We have integral of energy here because of time invariance. For confined systems, no integral of linear momentum because a translation transfers momentum to the system. What about angular momentum? Suppose a gas is contained in a sphere.

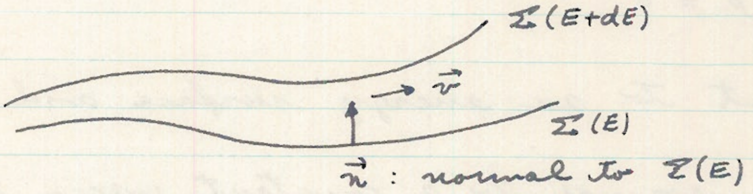


The sphere must have smooth walls to produce specular reflection, otherwise momentum is transferred to sphere producing torque. Then the angular momentum is not a constant of the motion. Hence we cannot make any statements about the linear and angular momentum in a confined gas, and cannot make statements about any of the rest of the C's because of details of walls, multivaluedness, etc.

Hence SM talks only about C_1 and ignores C_2 to C_{2F-1}.

Suppose we specify the energy. What does this do for ρ?

$E = \mathcal{H}(q_1 \cdots p_F)$, energy surfaces are denoted by $\Sigma(E)$ where Σ is surface not sum.



now: $\vec{n} \cdot \nabla \mathcal{H} = dE$; $\vec{n} \parallel \nabla \mathcal{H}$; $\therefore\ |\vec{n}| = \dfrac{dE}{|\nabla H|}$

also: $\nabla H = \left(\dfrac{\partial H}{\partial q_1}, \cdots, \dfrac{\partial H}{\partial p_F} \right)$

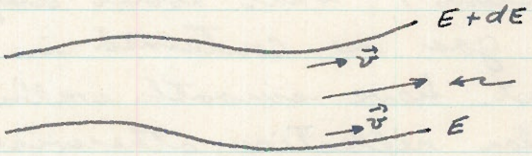
$$|\nabla H| = \left[\left(\frac{\partial H}{\partial q_1}\right)^2 + \cdots + \left(\frac{\partial H}{\partial p_F}\right)^2 \right]^{1/2}$$

$$= \left[(\dot{p}_1)^2 + \cdots + (\dot{q}_F)^2 \right]^{1/2} = \sqrt{\sum_j v_j^2} = |v| .$$

where $\vec{v}$ is the velocity in PS. Thus another way to write the normal is:

$$|\vec{n}| = \frac{dE}{|\vec{v}|}$$

Interpretation: Consider a hydrodynamic model, incompressible, $\nabla \cdot \vec{v} = 0$. $\vec{v}$ is flow vector in PS, tangent to a constant E surface.



← ρ does not vary in the direction $\perp$ or $\parallel$ to flow, $\vec{v}$.

$|\vec{v}| = \dfrac{dE}{|\vec{n}|}$ shows that $|\vec{v}| \downarrow$ as $|\vec{n}| \uparrow$

Since we work with energy, take for ρ:

$\rho(q_1 \cdots p_F) = C\, \delta(E - \mathcal{H}(q_1 \cdots p_F))$, zero unless at where we want it. C is normalization constant.

$C = \left\{ \int \delta(E-H)\, \Pi\, dq_j\, dp_j \right\}^{-1}$, then:

$$\rho(q_1 \cdots p_F) = \frac{\delta\left(E - H(q_1 \cdots p_F)\right)}{\int \delta\left(E - H(q_1 \cdots p_F)\right) \Pi\, dq_j\, dp_j}$$

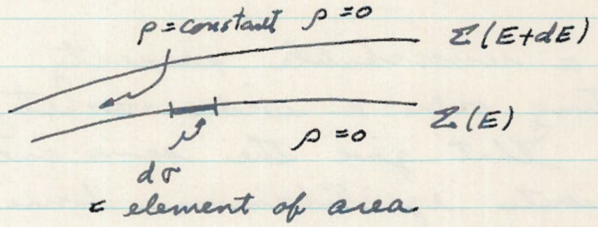
} We will use this a lot. Gibbs called it the microcanonical ensemble.

Boltzmann introduced method of finding ρ this way. We will call it the <u>microcanonical distribution function.</u>

LECTURE 5: 10-10-61

Recall the definition of the average value of a function of the PS. What can we say about ρ? It must be independent of the time for equilibrium. What other quantities are independent of time? Integrals of the motion. We know much only about one, the energy, and practically nothing about the rest, since they are very complicated and often multiple-valued.

Now one way of choosing ρ is:



The distances between surfaces is given by:

$$d\vec{n} = \frac{dE}{|\nabla H|} = \frac{dE}{|v|}$$

As the surfaces approach each other, $\rho \to \infty$ or a δ function on E.

An element of PS can now be written:

$$dV = \frac{d\sigma\, dE}{|\nabla H|}$$

If we consider E constant then the weight that attaches itself to the element of area $d\sigma$ is proportional to: $\frac{d\sigma}{|\nabla H|}$ or $\frac{d\sigma}{|v|}$

A more elegant way to put this is to say let:

$$\rho(q_1 \cdots p_F) = \frac{\delta[E - H(q_1 \cdots p_F)]}{\int \delta(E-H)\, dV}$$

which is the limiting value of the density as described above. For no good reason, this is called the microcanonical distribution function. We now take the microcanonical average of $\mathcal{G}(q_1 \cdots p_F)$:

$$\bar{\mathcal{G}} = \frac{\int \mathcal{G}(q_1 \cdots p_F)\, \delta(E-H)\, dV}{\int \delta(E-H)\, dV}$$

Or we can now write:

$$\bar{\mathcal{G}} = \frac{\int \bar{\mathcal{G}}(q_1 \cdots p_F)\, \delta(E-H)\, \frac{d\sigma\, dH}{|\nabla H|}}{\int \delta(E-H)\, \frac{d\sigma\, dH}{|\nabla H|}}$$

$$= \frac{\int_{\Sigma(E)} \mathcal{G}(q_1 \cdots p_F)\, \frac{d\sigma}{|\nabla H|}}{\int_{\Sigma(E)} \frac{d\sigma}{|\nabla H|}}$$

This result demonstrates a remarkable property. Even though we are integrating over a wide variety of velocities and positions that give the same energy, we say that $\bar{\mathcal{G}}$ represents equilibrium. Some of these states could hardly be called equilibrium; those where shock waves exist, or where a few molecules have tremendous velocities with the others at rest. How this makes sense is that these unusual events have only a very little weight on the constant E surface. Thus these events can be completely ignored because of their small weight. In SM, when things have small weight, they have *exceedingly* small weight.

Correspondence of Time and Ensemble Averages:

Historical Discussion: It was recognized by Boltzmann that it was possible to evaluate ensemble averages, but this says nothing about time averages since they presuppose a knowledge of the initial conditions. However, one measures time average. For many years it was considered the task of SM to prove the equivalence of time and ensemble averages.

We define the time average as follows:

$$(\mathcal{G})_{av} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathcal{G}(q_1(t) \cdots p_F(t), t)\, dt$$

This is the time average taken from some initial point $x_1(0) \cdots x_{2f}(0)$. Now in the 19th century it was assumed that one could only measure the time average and only calculate the ensemble average. for they maintained that they must prove the equivalence of time and ensemble averages. Glauber maintains that today we can measure ensemble averages. Because of improved instrumentation, we can now get instantaneous measurements, even of the position of gas molecules. This means we no longer have to prove the equivalence of time and ensemble averages. We will see that in Quantum Mechanics we will deal mainly with ensemble averages. Time averages are no more calculable in Quantum Mechanics than they are in Classical Mechanics. The extension to Quantum Mechanics would have eventually forced consideration of ensemble average anyway. However, we would still like to know what the correspondence is, if any. We will just sketch some of the more well-known attempts.

Assume that there is a time average. This is not a trivial statement. If it exists, what can we say about it? It is independent of the origin of time. Consider taking a different interval:

$$\frac{1}{T}\int_t^{T+t} G\,dt = \frac{1}{T}\left\{\int_0^{T+t} G\,dt - \int_0^t G\,dt\right\}$$

$$= \frac{1}{T}\,\frac{T+t}{T+t}\int_0^{T+t} G\,dt - \frac{1}{T}\int_0^t G\,dt$$

$$= \frac{1}{T+t}\int_0^{T+t} G\,dt + \frac{t}{T}\,\frac{1}{T+t}\int_0^{T+t} G\,dt - \frac{1}{T}\int_0^t G\,dt$$

Now if the time average exists, the last two terms go to 0 as $T \to \infty$:

$$\lim_{T\to\infty} \frac{1}{T}\int_t^{T+t} G\,dt = \lim_{T\to\infty} \frac{1}{T+t}\int_0^{T+t} G\,dt$$

Then if this limit exists, $(G)_{av}$ is constant along a trajectory because it is independent of the time origin. This lead Maxwell and Boltzmann to the equivalence of the time and ensemble averages. Suppose that the trajectory passes thru every point on the constant E surface if we wait long enough. This hypothesis of Maxwell and Boltzmann is called the:

Ergotic Hypothesis

from Ergon: energy and Hodon: Path.

<u>This hypothesis is absolutely false.</u> However, if it is true we are able to see that it leads to the identity of time and ensemble averages. Suppose one computes $\bar{G}$ and averages over time:

$$\bar{G} = (\bar{G})_{av}$$

We have just computed the time average of a constant so the statement is trivial. Suppose we do the inverse, first averaging over time at point x_0, and then over all x_0 on the constant E surface.

$$(\bar{G})_{av} = \overline{(G)_{av}}$$

Again we average over a constant. If the Ergotic Hypothesis were true, and all points on the path lie on the surface, we have implied.

$$\overline{(G)_{av}} = (G)_{av}$$

This is false because many dynamical systems have non-ergotic motion. The question of multiple-valuedness does not allow simple statements about ergoticity based on the integrals of motion to be made.

The main disproof was put forth by Rozenthal, Plancherel, 1913, who showed that a negligible amount of points lie on any one trajectory, thus completely destroying the Ergotic Hypothesis.

Ehrenfest brought forth the quasi-ergotic hypothesis. He maintained it was enough to have the trajectory come arbitrarily close to a given point. No progress has ever been made with this and it was never accepted.

There has been progress made in showing $\bar{G} = (G)av$ in a generalized dynamical system by Birkhoff, 1931. Consider the entire set of transformations that transform a space onto itself and preserves its measure: this is generalized dynamics. He showed that $(G)av$ exists if G is integrable over V, except for possible measure zero. If this is so, is $\bar{G} = (G)av$? Birkhoff assumed that V is metrically indecomposible, that is, it is not possible to separate V into two subspaces, each individually invariant to a transformation. Any such subsets must be of measure zero. $(G)av$ must then be the same for all points on the trajectory and on different trajectories. If this were not so, we could use this to decompose the space. If $(G)av$ is the same for all points, it is the same as $\bar{G}$. Metrical indecomposibility has never been proved.

Result: Today it is unclear where things stand regarding $\bar{G} = (G)av$. Interest has waned since we now realize that we almost always measure ensemble averages.

Calculation of Ensemble Averages:

How unique is the choice of the microcanonical distribution function? We will find that it is impossible to answer this because the observation of averages tells us very little about ρ. The averages are very insensitive to assumptions about ρ. Convenience rather than logic governs choice of ρ. We have not yet said why averages are insensitive to ρ. The reason is that we are not interested in very complicated things.

We almost always use one or two particle functions.

$$\sum_j f(q_j p_j)$$

$$\sum_{jk} g(q_j p_j, q_k p_k)$$

These are called sum functions and it is the sums that cause the insensitivity to ρ. They are almost constant all over E surface. This means we can use any weighting we please and still get the same average value. What we have chosen is a particularly simple one. The microcanonical distribution gives equal weighting to each volume element of PS. However, almost any will do. There is nothing logically dictated about ρ.

LECTURE 6: 10-17-61

Errata: $G(q_1(t) \cdots p_F(t), t)$ does not have to have an explicit dependence on time. If it does the time average does not exist.

Recall the microcanonical distribution function:

$$\rho = \frac{\delta(E - H(q_1 \cdots p_F))}{\int \delta(E-H)\, dV}$$

$$\bar{G} = \frac{\int \delta(E-H)\, G\, dV}{\int \delta(E-H)\, dV} = \frac{\int G \frac{d\sigma}{|\nabla H|}}{\int \frac{d\sigma}{|\nabla H|}} \qquad \text{Between } \Sigma(E) \text{ and } \Sigma(E+dE)$$

This says that all volumes of PS have equal a-priori probabilities. Averages are insensitive to the choice of ρ, and sum functions are nearly constant over a constant energy surface.

We can form some sort of analogy to this statement by considering a digital computer generating random numbers, N digits long, in binary form so there are 2^N possible numbers. If the numbers are truly random, the a priori probability of any given one is $1/2^N$.
The phase space consists of 2^N points and is discrete.
How can we check the impartiality of the choice of a given random number?

Label the digits of a number by a_j, $j = 1, \cdots, N$; a_j being either 1 or 0. Define the average of the digits in a number as:

$$A = \frac{1}{N} \sum_{j=1}^{N} a_j$$

Of course, this average will in general be very close to $1/2$.

Problem ②: Consider the deviations from $1/2$ and find the probability distribution around $1/2$. A way of doing this is to consider $N >> 1$ and change the Σ to $\int$. Then show the probability density about A is:

$$P(A) = \sqrt{\frac{2N}{\pi}}\, e^{-2N(A-\frac{1}{2})^2} \qquad \text{(Gaussian)}$$

in the asymptotic limit of N large and continuous. This could be done setting up the binomial coefficient, taking Stirling's approximation, and integrating. Also could be done using the central limit theorem.

Now consider $N = 10^{20}$: find to one significant figure the fraction of points in PS for which $|A-\frac{1}{2}| > 10^{-8}$, 10^{-9}, 10^{-10}.

This problem should show that when we sum on a random variable as is usual in physics, there is an indifference of A to the probability density or weighting given by the machine to different points in PS.

Definitions:

Name for the normalization integral in the microcanonical ensemble:

$$\text{Structure Function} = \Omega(E) = \int \delta\left(E - \mathcal{H}(q_1 \cdots p_F)\right) dV$$

Therefore: $$\rho(q_1 \cdots q_F) = \frac{1}{\Omega(E)}\, \delta(E - \mathcal{H})$$

Define the Characteristic Function:

$$\psi_E(\mathcal{H}) = \begin{cases} 1 & \mathcal{H} \leq E \quad \text{, inside energy surface} \\ 0 & \mathcal{H} > 0 \quad \text{, outside energy surface} \end{cases}$$

Now we can concoct various functional representations for $\psi_E(\mathcal{H})$, for example:

$$\psi_E(\mathcal{H}) = \int_0^E \delta(E' - \mathcal{H})\, dE' = \int_{\mathcal{H}}^{\infty} \delta(E - E')\, dE'$$

What about its derivative?

$$\frac{d\psi_E(\mathcal{H})}{dE} = \delta(E - \mathcal{H})$$

Then we can write:

$$\bar{\mathcal{G}} = \frac{1}{\Omega(E)} \int \mathcal{G}\, \delta(E - \mathcal{H})\, dV = \frac{1}{\Omega(E)} \frac{d}{dE} \int \mathcal{G}\, \psi_E(\mathcal{H})\, dV$$

$$= \frac{1}{\Omega(E)} \frac{d}{dE} \int_{\mathcal{H} \leq E} \mathcal{G}\, dV$$

The special case of $\mathcal{G} = 1$ gives us $\Omega(E)$:

$$\Omega(E) = \frac{d}{dE} \int_{\mathcal{H} \leq E} dV$$

Equipartition Theorems:

We assume nothing more than that $\rho(\mathcal{H})$ depends on the PS coordinates only through the Hamiltonian (stationary). Also assume $\int \rho(\mathcal{H})\, dV = 1$. This tells us that ρ must decrease very fast to cancel out the rise in $\mathcal{H}$ as the coordinates become very large. We can always assume ρ decreasing exponentially is sufficient to do this.

Define the $\underline{\Theta \text{ function}}$:

$$\theta(E) = \int_E^\infty \rho(E')\, dE'$$

If ρ is exponentiol, than it follows that θ must be also.

Now consider a function of the coordinates, in particular consider the average value of $p_j \frac{\partial H}{\partial p_j}$ to show the use of $\theta(E)$:

$$\overline{p_j \frac{\partial \mathcal{H}}{\partial p_j}} = \int \rho(\mathcal{H})\, p_j \frac{\partial H}{\partial p_j}\, dV \quad ; \text{ no sum on } j.$$

now:

$$\frac{d\theta(E)}{dE} = -\rho(E) \quad , \text{ therefore:}$$

$$\overline{p_j \frac{\partial \mathcal{H}}{\partial p_j}} = -\int p_j \frac{\partial \theta(\mathcal{H})}{\partial \mathcal{H}} \frac{\partial \mathcal{H}}{\partial p_j}\, dV$$

$$= -\int p_j \frac{\partial \theta(\mathcal{H})}{\partial p_j} \prod_k dq_k\, dp_k$$

We can write this in the form:

$$\overline{p_j \frac{\partial \mathcal{H}}{\partial p_j}} = \underbrace{-\int \frac{\partial}{\partial p_j}\left(p_j\, \theta(\mathcal{H})\right) \prod_k dq_k\, dp_k}_{0} + \int \theta(\mathcal{H})\, dV$$

because $\theta(E) \Rrightarrow 0$ rapidly as $E \to \infty$

Then for all j and k:

$$\overline{p_j \frac{\partial \mathcal{H}}{\partial p_j}} = \overline{p_k \frac{\partial \mathcal{H}}{\partial p_k}} = \int \theta(\mathcal{H})\, dV$$

This idea works just as well for the space coordinates providing we put the system in a potential box so that we can say the potential is infinite beyond some coordinate and $\theta(\mathcal{H}) \Rrightarrow 0$ also. We get:

$$\overline{q_j \frac{\partial \mathcal{H}}{\partial q_j}} = -\int q_j \frac{\partial \theta}{\partial q_j}\, dV = \int \theta(\mathcal{H})\, dV$$

We now calculate θ explicitly for the microcanonical distribution function:

$$\rho(E') = \frac{\delta(E-E')}{\Omega(E)}$$

$$\theta(\mathcal{H}) = \frac{1}{\Omega(E)} \int_{\mathcal{H}}^{\infty} \delta(E-E')\, dE = \frac{1}{\Omega(E)} \psi_E(\mathcal{H})$$

$\theta(\mathcal{H})$ will appear all the time in equipartition problems. Now:

$$\int \theta(\mathcal{H})\, dV = \frac{1}{\Omega(E)} \int \psi_E(\mathcal{H})\, dV = \frac{1}{\Omega(E)} \int_{\mathcal{H}<E} dV = \frac{V(E)}{\Omega(E)}$$

But: $\Omega(E) = \dfrac{d}{dE} \displaystyle\int_{\mathcal{H}<E} dV = \dfrac{d}{dE} V(E)$

Therefore:

$$\int \theta(\mathcal{H})\, dV = \frac{1}{\frac{1}{V(E)} \frac{d}{dE} V(E)} = \frac{1}{\frac{d}{dE} \log V(E)}$$

Then, in the microcanonical ensemble:

$$\overline{p_j \frac{\partial \mathcal{H}}{\partial p_j}} = \overline{q_j \frac{\partial \mathcal{H}}{\partial q_j}} = \int \theta(\mathcal{H})\, dV = \frac{1}{\frac{d}{dE} \log V(E)}$$

or: $$\overline{p_j \dot{q}_j} = -\overline{q_k \dot{p}_k} = \int \theta(\mathcal{H})\, dV = \frac{1}{\frac{d}{dE} \log V(E)}$$

We now make some further assumptions about $\mathcal{H}$.

Assume that the kinetic energy is separable into sums of the kinetic energy of some subsystems. That is, assume we can write:

$$\mathcal{H} = \sum_{\alpha} T_\alpha (p_1^\alpha \cdots p_{f\alpha}^\alpha , q_1 \cdots) + V(q_1 \cdots q_F)$$

where $f\alpha$ is the number of momenta of subsystem α. This Hamiltonian covers many physical problems except those of the EM field. We can further assume that T is a very general quadratic in the momenta. That is, we assume that T_α's are homogeneous in the second degree. To show this, suppose each momentum contained the same constant parameter λ. Then we have:

$$T_\alpha (\lambda p_1^\alpha \cdots \lambda p_{f\alpha}^\alpha , q_1 \cdots) = \lambda^2 T_\alpha (p_1^\alpha \cdots p_{f\alpha}^\alpha , q_1 \cdots)$$

Differentiate by $\frac{d}{d\lambda}$ both sides:

$$\sum_j p_j \frac{\partial}{\partial(\lambda p_j)} T_\alpha (\lambda p_1^\alpha \cdots) = 2 \lambda T_\alpha (p_1 \cdots)$$

Now let $\lambda \to 1$:

$$\sum_j p_j \frac{\partial}{\partial p_j} T_\alpha = 2 T_\alpha \qquad \text{(Euler's Theorem)}$$

Now take the average:

$$\overline{T_\alpha} = \frac{1}{2} \sum_j \overline{p_j \frac{\partial \mathcal{H}}{\partial p_j}} = \frac{f_\alpha}{2} \int \theta(\mathcal{H})\, dV = \frac{f_\alpha}{2} \frac{1}{\frac{d}{dE} \log V(E)}$$

Then the average kinetic energy of such a subsystem is just the number of degrees of freedom times $\int \theta(\mathcal{H})\, dV$.

This is the most general statement of the equipartition theorem.

now extend to include all types and varieties of harmonic oscillators, coupled together in any possible way, assuming the q's cartesian and the kinetic energy independent of the q's. Now V has a homogeneous quadratic form in $q_1 \cdots q_N$ and we have the statement:

$$\sum_{j=1}^{N} q_j \frac{\partial V}{\partial q_j} = 2V$$

and $$\sum_{j=1}^{N} \overline{q_j \frac{\partial V}{\partial q_j}} = \sum_{j=1}^{N} \overline{q_j \frac{\partial H}{\partial q_j}} = 2\overline{V} = N \int \theta(\mathcal{H})\, dV$$

Therefore: $$\overline{V} = \frac{N}{2} \int \theta(\mathcal{H})\, dV = \overline{\mathcal{K}}$$

and for a system of harmonic oscillators, the average potential and kinetic energies are equal.

LECTURE 7 : 10-19-61

Recall from last lecture:

$$\overline{p_j \frac{\partial H}{\partial p_j}} = \overline{q_k \frac{\partial H}{\partial q_k}} = \int \theta(H)\, dV$$

When we deal with separate subunits, we can say:

$$\mathcal{K}_\alpha = \frac{f_\alpha}{2} \int \theta(\mathcal{H})\, dV$$

Consider an example of the type of $\mathcal{K}$ that is not separable:

$$f_1(q)\, p_1^2 + f_{12}\, p_1 p_2 + f_2\, p_2^2$$

This kinetic energy term has two degrees of freedom but cannot be separated because of the cross-term.

Consider now a gas of particles defined by their position and momenta:

$$\vec{r}_\alpha \qquad p_j^\alpha \; ; \; j = 1, 2, 3$$

α labels the particular molecule. We consider each molecule a sub-system: we obtain:

$$\sum_{j=1}^{3} \overline{p_j^\alpha \frac{\partial H}{\partial p_j^\alpha}} = 2\,\overline{\mathcal{T}_\alpha} = 3 \int \theta(H)\, dV$$

$$= \sum_{j=1}^{3} \overline{q_j^\alpha \frac{\partial H}{\partial q_j^\alpha}} = \overline{\vec{r}^\alpha \cdot \nabla^\alpha V}$$

We define: $\vec{F} = -\nabla V$, now recall from kinetic theory the Virial Theorem and we see that this is what we have shown above:

$$\overline{\vec{r}_\alpha \cdot \vec{F}^\alpha} = -2\,\overline{\mathcal{T}_\alpha}$$

This was first demonstrated by Clausius in 1870 and we now offer his method:

$$m \ddot{\vec{r}}_\alpha = \vec{F}^\alpha$$

Definition of the Virial: $W = (\vec{r}_\alpha \cdot \vec{F}^\alpha)_{\text{time average}}$

$$= \frac{1}{T} \int_0^T \vec{r}^\alpha \cdot \vec{F}^\alpha \, dt = \frac{1}{T} \int_0^T m\, \vec{r}^\alpha \cdot \ddot{\vec{r}}^\alpha \, dt$$

$$= \frac{1}{T} \int_0^T m \left\{ \frac{d}{dt} (\vec{r}^\alpha \cdot \dot{\vec{r}}^\alpha) - (\dot{r}_\alpha)^2 \right\} dt$$

$$= \frac{1}{2T} \int_0^T m \frac{d^2}{dt^2} (r^\alpha)^2 \, dt - 2 (\mathcal{T}_\alpha)_{ta}$$

or:

$$W = \frac{1}{2T} m \left\{ (\vec{r}^\alpha \cdot \dot{\vec{r}}^\alpha)_T - (\vec{r}^\alpha \cdot \dot{\vec{r}}^\alpha)_0 \right\} - 2 (\mathcal{T}_\alpha)_{ta}$$

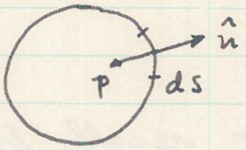
now as long as the system is bounded in r_α, the { } must remain finite so that this term vanishes as $T \to \infty$.

Hence, in this limit we have:

$$(\vec{r}^\alpha \cdot \vec{F}^\alpha)_{ta} = -2\,(T_\alpha)_{ta}$$

The Virial theorem originally found great application in problems involving equations of state of gases.

What is the virial for a perfect gas? Here there are no forces between the gas and container except impulsive forces:



$$\vec{F} = -p\hat{n}$$

Then: $$\sum_\alpha \overline{\vec{r}^\alpha \cdot \vec{F}^\alpha} = -\int_S p\,\hat{n}\cdot\vec{r}\,dS$$

$$= -p\int_S \hat{n}\cdot\vec{r}\,dS = -p\int_V \underbrace{(\nabla\cdot\vec{r})}_{3}\,d\tau = -3pv$$

$$= -F\int \theta(H)\,dV = -2\sum_\alpha \overline{(T_\alpha)}$$

where v is the volume of the container and $F = 3N$, the number of degrees of freedom.

Now:

$\sum_\alpha \overline{T_\alpha}$ is the total energy of the gas since there is no interaction between particles, hence:

$$\sum_\alpha \overline{T_\alpha} = E$$

Then: $$pv = \frac{2}{3}E = N\int \theta(H)\,dV$$

From this relation we can get a sneak preview of the meaning of temperature. Recall the perfect gas law:

$$pv = NkT$$

We can identify $E = \frac{3}{2}NkT = \frac{3}{2}N\int\theta(H)\,dV$

Therefore: $kT = \int \theta(H)\, dV$. Or, in the microcanonical distribution:

$$kT = \frac{1}{\frac{d}{dE} \log V(E)}$$

We take this as the definition of temperature for the time being.

Computation of Distribution Functions

We consider the gas in a more general way now. Take as the model N particles, $F = 3N$ degrees of freedom, $6N$ coordinates, all masses the same, and no interaction.

$$H = \frac{1}{2m} \sum_{j=1}^{F} p_j^2$$, for q_j's inside a leak-proof container.

Then:

$$\rho = \frac{\delta\left(E - \frac{1}{2m} \sum_{j=1}^{F} p_j^2\right)}{\Omega(E)}$$

$$\Omega(E) = \int \delta\left(E - \frac{1}{2m} \sum_{j=1}^{F} p_j^2\right) \prod_{j=1}^{F} dq_j\, dp_j$$

$$= v^N \int \delta\left(E - \frac{1}{2m} \sum p_j^2\right) \prod dp_j$$

Suppose we are interested in the distribution of one component of momentum of one particle, throwing away the rest by integrating over them:

$$P(p_1) = \int \rho(p_1 \cdots p_F, q_1 \cdots q_F) \prod_{j=2}^{F} dp_j \prod_{j=1}^{F} dq_j$$

We must satisfy the condition for normality:

$$\int P(p_1)\, dp_1 = 1$$, hence:

$$P(p_1) = \frac{\int \delta\left(E - \frac{p_1}{2m} - \frac{1}{2m} \sum_{j=2}^{F} p_j^2\right) \prod_{k=2}^{F} dp_k}{\int \delta\left(E - \frac{1}{2m} \sum_{j=1}^{F} p_j^2\right) \prod_{k=1}^{F} dp_k}$$

The v^N's cancel. We will show that this eventually leads to the Maxwell-Boltzmann (MB) distribution. Integration goes over an n-dimensional sphere in the kind of integrations indicated above. We hence digress for a moment to consider some properties of n-dimensional spheres, particularly their volume and surface.

Call the surface area of the unit sphere A_n. Then obviously for radius r, the area is:

$$A_n r^{n-1}$$

Consider the integral:

$$\int_{-\infty}^{\infty} e^{-\sum_{j=1}^{n} x_j^2} \prod_{k=1}^{n} dx_k = \left[\int_{-\infty}^{\infty} e^{-x^2} dx\right]^n = \pi^{n/2}$$

Now let $r^2 = \sum_{j=1}^{n} x_j^2$ and we get for the above, noting that:

$$\prod_{k=1}^{n} dx_k = A_n r^{n-1} dr,$$

$$A_n \int_0^{\infty} e^{-r^2} r^{n-1} dr = \pi^{n/2}$$

Now let $s = r^2$, $ds = 2r\,dr$, $r^{n-1} = s^{\frac{n-1}{2}}$, and get:

$$\frac{1}{2} A_n \int_0^{\infty} s^{\frac{n}{2}-1} e^{s} ds = \pi^{n/2}$$

But $\int_0^{\infty} s^{\frac{n}{2}-1} e^{s} ds \equiv \Gamma(\frac{n}{2})$

Some properties of the Γ function are:

$\Gamma(n+1) = n!$ for n an integer

Otherwise: $\Gamma(n+1) = n\Gamma(n)$; and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

We now have:

$$\boxed{A_n = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}}$$

Some of the more useful A_n's are:

$A_0 = 0$
$A_1 = 2$
$A_3 = 2\pi$

$A_3 = 4\pi$
$A_4 = 2\pi^2$
$A_5 = \frac{8}{3}\pi^2$

$A_6 = \pi^3$

Now that we have an expression for A_n, we consider the volumes of hyperspheres:

$$V = \int_0^r A_n r^{n-1}\, dr = \frac{A_n}{n} r^n$$

$$\frac{A_n}{n} = \frac{2}{n} \frac{\pi^{n/2}}{\Gamma(\frac{n}{2})} = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$$

Hence : $$\boxed{V = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r^n}$$

Now back to evaluating distribution functions. The kind of integral we want to do are of the form:

$$\mathcal{I} = \int_{-\infty}^{\infty} \delta(r^2 - \sum x_j^2) \prod dx_j$$

Define the radial coordinates: $X^2 = \sum_{j=1}^{n} x_j^2$
Then:

$$\mathcal{I} = \int_0^{\infty} \delta(r^2 - X^2)\, A_n X^{n-1} dX$$

Recall some of the properties of δ functions. What is $\delta[f(x)]$? The δ function acts when $f(x) = 0$ giving the roots x_0. Expand $f(x)$ about one of these roots:

$$f(x) \simeq \underbrace{f(x_0)}_{0} + (x - x_0) f'(x_0)$$

Then: $$\delta\{f(x)\} = \delta\left[(x-x_0) f'(x_0)\right] = \frac{1}{|f'(x_0)|} \delta(x - x_0),$$

From the property that $\delta(ax) = \frac{1}{|a|}\delta(x)$ as can be shown from a representation of the δ function.

Then, $\delta(r^2 - x^2) = \frac{1}{2r}\delta(x \pm r)$, Then:

$$\boxed{I = \int_0^\infty \delta(r^2 - x^2)\, A_n x^{n-1}\, dx = \frac{1}{2} r^{n-2} A_n}$$

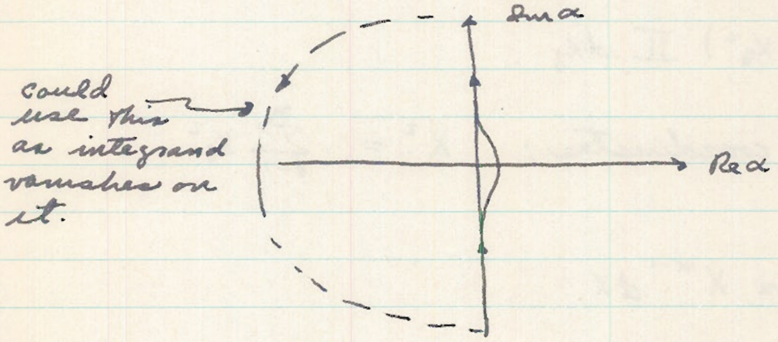
We now digress to do the same thing by another technique that will be used later on. This is the method of representing the δ function by an integral. We choose this representation to be the following:

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\lambda x}\, d\lambda$$

We now let $i\lambda = \alpha$ and get:

$$\delta(x) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{\alpha x}\, d\alpha$$

Usually the contour would the imaginary axis and the left-half plane, but since $e^{\alpha x}$ is analytic everywhere in the finite plane, we could deform the contour into, for example:



Now we can write for I:

$$I = \frac{1}{2\pi i}\int_{-\infty}^{\infty}\int_{-i\infty}^{i\infty} e^{\alpha\left(r^2 - \sum_{j=1}^{n} x_j^2\right)}\, d\alpha \prod_{j=1}^{n} dx_j$$

$$= \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} d\alpha\, e^{\alpha r^2}\left[\int_{-\infty}^{\infty} e^{-\alpha x^2}\, dx\right]^n$$

Consider: $\int_{-\infty}^{\infty} e^{-\alpha x^2} dx$. This only has a meaningful result for $\mathrm{Re}\,\alpha > 0$ as it diverges for $\mathrm{Re}\,\alpha < 0$.

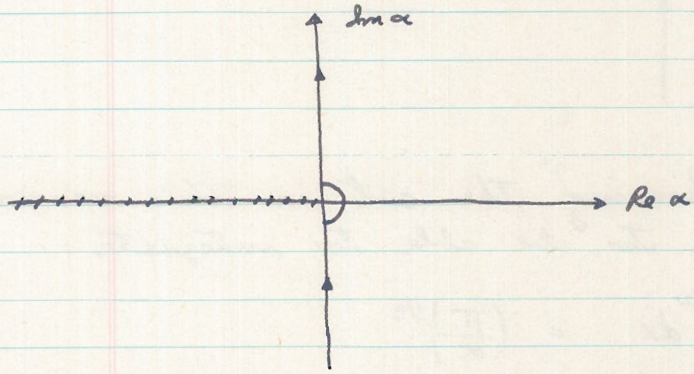
This is the reason for deforming the contour into the right half plane (RHP). Here an explicit evaluation is possible and we have:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

Since we have a multivalued function in $1/\alpha^{1/2}$ we must decide where to put a branch cut. Do this along the negative real axis since we can only evaluate the integral for $\mathrm{Re}\,\alpha > 0$ and we know it is analytic there, and then continue it into the LHP, except along $\mathrm{Re}\,\alpha < 0$, by analytic continuation. We can then write:

$$f = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \left(\frac{\pi}{\alpha}\right)^{n/2} e^{\alpha r^2} d\alpha$$

The contour runs as shown:



LECTURE 8 : 10-24-61

Recapitulation :

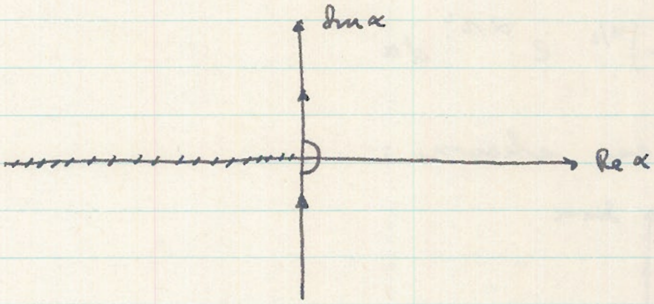
$$\mathcal{I} = \int_{-\infty}^{\infty} \delta\left(r^2 - \sum_j x_j^2\right) \prod dx_j = \tfrac{1}{2} A_n r^{n-2}$$

where $A_n = \dfrac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}$

$$\delta(x) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{\alpha x} \, d\alpha$$

$$\mathcal{I} = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{\alpha r^2} \left(\frac{\pi}{\alpha}\right)^{n/2} d\alpha$$

where the contour is:

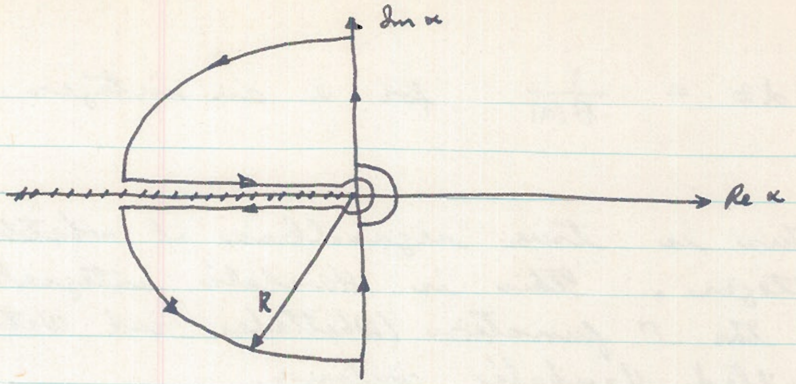


The reason for choosing the cut in this manner is because we want to be able to integrate:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} \, dx = \left(\frac{\pi}{\alpha}\right)^{1/2}$$

which we can do only for Re α > 0. However, we can make an analytic continuation into the LHP providing we place the branch cut where it is.

For the evaluation of $\mathcal{I}$, then, consider the evaluation of its integrand along the following contour:



For Re α < 0, n > 0, it is easily seen that the integrand of $\mathscr{I}$ vanishes along the quarter-circles as $R \to \infty$. Since there are no poles within the contour, we may write:

$$\frac{1}{2\pi i} \int_{\uparrow} e^{\alpha r^2} \left(\frac{\pi}{\alpha}\right)^{n/2} d\alpha + \frac{1}{2\pi i} \int_{\rightleftarrows\circ} e^{\alpha r^2} \left(\frac{\pi}{\alpha}\right)^{n/2} d\alpha = 0$$

or:

$$\mathscr{I} = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{\alpha r^2} \left(\frac{\pi}{\alpha}\right)^{n/2} d\alpha = \frac{1}{2\pi i} \int_{\leftrightarrows\circ} e^{\alpha r^2} \left(\frac{\pi}{\alpha}\right)^{n/2} d\alpha$$

Suppose we let: $z = \alpha r^2$, then:

$$\mathscr{I} = \pi^{n/2} r^{n-2} \frac{1}{2\pi i} \int_{\leftrightarrows\circ} \frac{e^z}{z^{n/2}} dz$$

Consider now the integral: $\frac{1}{2\pi i} \int_{\leftrightarrows\circ} \frac{e^z}{z^{n/2}} dz = \frac{1}{2\pi i} \int_{\leftrightarrows\circ} \frac{e^z}{z^a} dz$

with $n/2 = a =$ an integer. Hence there are no branches because the integrand is uniquely defined over the z plane. The contour reduces to a circle about the origin, and we can use the residue theorem. Expand the integrand in a Laurent series:

$$\frac{e^z}{z^a} = \sum_{l=0}^{\infty} \frac{z^{l-a}}{l!}$$

The coefficient of z^{-1} is $\frac{1}{(a-1)!} = \frac{1}{\Gamma(a)}$

Then: $\frac{1}{2\pi i}\int_{\bigcirc} \frac{e^z}{z^a}\, dz = \frac{1}{\Gamma(a)}$ for a an integer.

More generally, this is true regardless of whether or not a is an integer. This is Hankel's integral representation of the Γ function (Whittaker and Watson). The contour is called Hankel's contour.

$$\frac{1}{2\pi i}\int_{\circlearrowleft} \frac{e^z}{z^a}\, dz = \frac{1}{\Gamma(a)}$$

Thus, we obtain the same result as previously:

$$\boxed{I = \frac{\pi^{n/2}\, r^{n-2}}{\Gamma(\frac{n}{2})}}$$

We will now proceed to use these in the calculation of $P(p_1)$. Recall:

$$P(p_1) = \frac{\int \delta\left(E - \frac{p_1^2}{2m} - \sum_{j=2}^{F}\frac{p_j^2}{2m}\right)\prod_{j=2}^{F} dp_j}{\int \delta\left(E - \sum_{j=1}^{F}\frac{p_j^2}{2m}\right)\prod_{j=1}^{F} dp_j}$$

$$= \frac{\int \delta\left(2mE - p_1^2 - \sum_{j=2}^{F} p_j^2\right)\prod_{j=2}^{F} dp_j}{\int \delta\left(2mE - \sum_{j=1}^{F} p_j^2\right)\prod_{j=1}^{F} dp_j} \qquad \left.\begin{matrix}(n = F-1)\\ \\ (n = F)\end{matrix}\right\}\ \text{number of dimensions}$$

$$= \frac{\pi^{\frac{F-1}{2}}}{\Gamma(\frac{F-1}{2})}\left(2mE - p_1^2\right)^{\frac{F-3}{2}} \cdot \frac{\Gamma(\frac{F}{2})}{\pi^{F/2}\,(2mE)^{\frac{F-2}{2}}}$$

Then:

$$\boxed{P(p_1) = \frac{1}{\sqrt{\pi}}\,\frac{\Gamma(\frac{F}{2})}{\Gamma(\frac{F-1}{2})}\,\frac{1}{\sqrt{2mE}}\left(1 - \frac{p_1^2}{2mE}\right)^{\frac{F-3}{2}}}$$

Now this is the precise result for the microcanonical distribution, <u>not</u> the MB as is usually seen.

However, we do have very often the situation where $\frac{p_1^2}{2mE}$ is very small and F is very large. Recall from elementary calculus the definition of the exponential:

$$e = \lim_{x \to 0} (1 + x)^{1/x}$$

What we have here is similar. Write:

$$(1-x)^n = \left(e^{\ln(1-x)} \right)^n = e^{n \log(1-x)}$$

Note that above:

$$\lim_{x \to 0} (1+x)^{1/x} = e^{\lim_{x \to 0} \frac{\ln(1+x)}{x}} = e$$

so we are really taking $\lim_{x \to 0} \frac{\ln(1+x)}{x}$. Continuing:

Now expand $\ln(1-x)$ and get,

$$(1-x)^n = e^{-n\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots\right)} = e^{-nx} \, e^{-n\frac{x^2}{2}} \cdots$$

It is possible for this product to converge very rapidly providing:

$$\frac{n x^2}{2} << 1 \quad ; \text{ or } \quad x << \sqrt{\frac{2}{n}}$$

Then we have:

$$(1-x)^n \approx e^{-nx}$$

which is the elementary calculus limit. Note that we are able to evaluate the limitations, however.

Then we can write:

$$P(p_1) = \frac{\Gamma\left(\frac{F}{2}\right)}{\sqrt{\pi} \, \Gamma\left(\frac{F-1}{2}\right)} \frac{1}{\sqrt{2mE}} \, e^{-\frac{p_1^2}{2mE} \frac{F}{2}}$$

where we have taken: $x = \frac{p_1^2}{2mE}$; $n = \frac{F-3}{2} \approx \frac{F}{2}$

since $F >> 1$.

now that we have made the above approximation, it would be good to renormalize. However, let's see how the Γ functions behave first:

$$\frac{\Gamma(n+1)}{\Gamma(n)} = n \quad ; \quad \text{can we say} \quad \frac{\Gamma(n+1/2)}{\Gamma(n)} \simeq \sqrt{n} \ ?$$

Consider:

$$\log \frac{\Gamma(n+1/2)}{\Gamma(n)} = \log \Gamma(n+1/2) - \log \Gamma(n)$$

$$\simeq \frac{1}{2} \frac{d}{dn} \log \Gamma(n) \simeq \frac{1}{2} \left\{ \log \Gamma(n+1) - \log \Gamma(n) \right\}$$

$$= \frac{1}{2} \log n \quad , \text{since} \quad \Gamma(n+1) = n\Gamma(n)$$

Therefore: $\quad \dfrac{\Gamma(n+1/2)}{\Gamma(n)} \simeq \sqrt{n}$

The argument is ambiguous but it does motivate the replacement for large n. The error in the above replacement is essentially in the high order terms of a power series expansion of $\log \Gamma(n+1/2)$:

$$\log \Gamma(n+1/2) = \log \Gamma(n) + \frac{1}{2} \frac{d}{dn} \log \Gamma(n) + \frac{1}{8} \frac{d^2}{dn^2} \log \Gamma(n) + \cdots$$

$$\frac{d}{dn} \log \Gamma(n) = \frac{d}{dn} \log (n-1)! \simeq \frac{d}{dn} (n \ln n - n)$$

$$\simeq \log n$$

$$\frac{d^2}{dn^2} \log \Gamma(n) \simeq \frac{1}{n}$$

so the error involved is of the order of $\frac{1}{n}$, a very small number for many degrees of freedom.

$$\boxed{\rho(p_1) = \left[\frac{F}{2\pi \cdot 2mE} \right]^{1/2} e^{-\frac{p_1^2}{2m} \frac{F}{2E}} \quad ; \quad F = 3N}$$

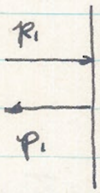
taking:

$$n = \frac{F}{2}.$$

We may also write:

$$P(p_1) = \sqrt{\frac{1}{2\pi m}\;\frac{3N}{2E}}\; e^{-\frac{p_1^2}{2m}\left(\frac{3N}{2E}\right)}$$

We will now obtain the equation of state by some quantitative arguments. Take as the model a smoothed wall container and to find the pressure, consider one component of p normal to the wall.



$$P = \frac{N}{V}\int_{p_1>0} \frac{p_1}{m}\; 2p_1\; P(p_1)\, dp_1$$

where $\frac{N}{V}$ is the density of particles, $\frac{p_1}{m}$ the velocity, and $2p_1 P(p_1)$ the distribution in the direction normal to the wall. Then:

$$P = \frac{N}{V}\int_{-\infty}^{\infty} \frac{p_1^2}{m}\, P(p_1)\, dp_1 = \frac{N}{V}\,\frac{\overline{p_1^2}}{m}$$

Recall some properties of the Gaussian distribution, which is what $P(p_1)$ above is.

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{x^2}{2\sigma^2}} \quad ; \quad \text{with } \overline{x^2} = \sigma^2$$

Then, making the proper identification:

$$P = \frac{N}{V}\,\frac{1}{m}\,\frac{2mE}{3N} = \frac{N}{V}\,\frac{2E}{3N} \quad ; \quad PV = N\,\frac{2E}{3N}$$

By recourse to the perfect gas law, we can identify in an off-hand manner:

$$kT \to \frac{2E}{3N}$$

Now we see we have the usual form of the MB distribution:

$$\boxed{P(p_1) = \frac{1}{\sqrt{2\pi m kT}}\, e^{-\frac{p_1^2}{2mkT}}}$$

We can ask how good is the MB approximation? Generally it is good as long as p_1 is small compared with $2mE$ and F is large compared to one.

Recall the condition:

$$x << \sqrt{\frac{2}{n}}$$

or:
$$\frac{p_1^2}{2mE} << \sqrt{\frac{4}{F}} = \sqrt{\frac{4}{3N}}$$

or:
$$\frac{p_1^2}{2m} << \sqrt{3N}\left(\frac{2E}{3N}\right) = \sqrt{3N}\, kT$$

so the MB distribution is a poor approximation unless:

$$\frac{p_1^2}{2m} << \sqrt{3N}\, kT$$

Now at 300°K: $kT \sim \frac{1}{40}$ ev

We take $N = 10^{22}$ with the result that:

$$\sqrt{3N}\, kT \approx 1.7 \cdot 10^{11} \times \frac{1}{40} = 4 \cdot 10^{9}\,\text{ev}$$

Now it so happens that a thermal molecule seldom attains energies of 4 BEV, so the MB distribution is practically always an excellent approximation.

LECTURE 9 : 10-26-61

Recall the derivation of MB from the microcanonical distribution for a non-interacting system of free particles, considering only one component of momentum. Now ask for three components of the momentum: $P(p_1 p_2 p_3)$. The calculation is essentially the same as before using $\Omega(E)$ over all phase space in the denominator and using an integration over phase space minus 3 dimensions. The limiting result is again a MB distribution:

$$P(p_1 p_2 p_3) = \left[\frac{1}{2\pi m}\left(\frac{3N}{2E}\right)\right]^{3/2} e^{-\frac{p^2}{2m}\left(\frac{3N}{2E}\right)}$$

where $p^2 = p_1^2 + p_2^2 + p_3^2$.

We should note that before going to the asymptotic limit, $P(p_1 p_2 p_3)$ does not have the separability in the asymptotic limit into 3 identical momentum distribution functions that the MB form does.

We have been considering surfaces of hyperspheres. Now let's talk about the volume of F-dimensional hyperspheres:

$$V_F(r) = \frac{2\pi^{F/2}}{F\,\Gamma\left(\frac{F}{2}\right)}\, r^F = \frac{\pi^{F/2}}{\Gamma\left(\frac{F}{2}+1\right)}\, r^F$$

The rapid convergence of V_F for large F will mean that most of the volume of the hypersphere is near the surface with almost none in the middle. To get an idea of this rapid convergence, subtract a slightly smaller hypersphere from $V_F(r)$:

$$V_F(r) - V_F(r-t) = \frac{\pi^{F/2}}{\Gamma\left(\frac{F}{2}+1\right)}\left\{ r^F - (r-t)^F \right\}$$

$$= V_F(r)\left\{1 - \left(1-\frac{t}{r}\right)^F\right\}$$

We can proceed as before, taking $\frac{t}{r} \ll \sqrt{\frac{2}{F}}$, which invokes the limiting case of:

$$\left(1 - \frac{t}{r}\right)^F \doteqdot e^{-t/r\,F}$$

Then: $V_F(r) - V_F(r-t) = V_F(r)\left[1 - e^{-\frac{t}{r}F}\right]$

For $F = 10^{22}$, the approximation requires that $\frac{t}{r} \ll 10^{-11}$. Consider $\frac{t}{r} = 10^{-22}$ or only a very small difference between the radii of the two spheres; then:

$$V_F(r) - V_F(r-t) = V(r)\left\{1 - e^{-1}\right\}$$

Thus all but 1/3 of V_F lies in a shell $10^{-22}r$ thick from the surface, and hence a great deal of the volume lies in an infinitesimal surface shell. This essentially small difference between whether we consider the volume or surface of the hypersphere shows how sloppy one can be in some considerations in statistical mechanics.

Problem ③:

Assume a perfect gas with $F = 3N$ momentum components. Choose as ρ a function not larger than a certain energy.

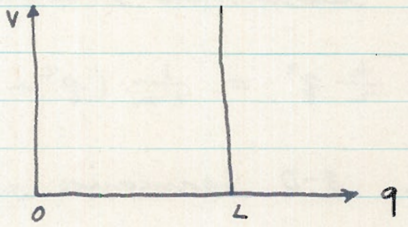
$$\rho\,(q_1 \cdots p_1 \cdots) = \text{constant} \times \psi_E(\mathcal{H}) \; ; \quad \psi_E(\mathcal{H}) = \begin{cases} 0 & \mathcal{H} > E \\ 1 & \mathcal{H} \leq E \end{cases}$$

One finds as a result for $P(p_1)$ the MB distribution because of the convergence of the volume of a hypersphere into its surface. Because of this, the energy of the system is practically E even though $\psi_E(\mathcal{H})$ does not tell us this.

We now consider some computational examples of the structure function $\Omega(E)$:

Particle in a 1-D Box:

Consider:



$$\Omega(E) = \int \delta\left(E - \frac{p^2}{2m}\right) dp\, dq$$

$$= L \cdot 2 \cdot 2m \cdot \frac{1}{2|p|} = L\sqrt{\frac{2m}{E}}$$

using the fact that:

$$\delta\left(E - \frac{p^2}{2m}\right) = 2m\, \delta\left(2mE - p^2\right)$$

$$= 2m \cdot \frac{1}{2\sqrt{2mE}}\, \delta\left(p \pm \sqrt{2mE}\right)$$

Particle in an n-dimensional hypercube:

$$\Omega(E) = \int \delta\left(E - \frac{1}{2m}\sum_{j=1}^{n} p_j^2\right) \prod_{j=1}^{n} dp_j\, dq_j$$

$$= L^n \cdot 2m \int \delta\left(2mE - \sum_{j=1}^{n} p_j^2\right) \prod_{j=1}^{n} dp_j$$

Use the formula: $\int_0^\infty \delta(r^2 - x^2)\, A_n\, x^{n-1}\, dx = \frac{1}{2} r^{n-2} A_n$

Then:

$$\Omega(E) = L^n \cdot 2m \cdot \frac{1}{2} A_n\, (2mE)^{\frac{n-2}{2}}$$

or: $$\Omega(E) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)} L^n\, 2m\, (2mE)^{\frac{n-2}{2}}$$

Note: in 1-D, we have Ω(E) ↓ as E ↑; in 2-D there is no change; for higher dimensions, we have Ω(E) ↑ as E ↑.

Problem ④ :

(a) For the 1-D harmonic oscillator, find $\Omega(E)$:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{k}{2} q^2 = \frac{1}{2m}\left(p^2 + m^2 \omega^2 q^2\right) \quad ; \quad k = m\omega^2$$

(b) Find $\Omega(E)$ for the 3-D harmonic oscillator:

$$\mathcal{H} = \frac{1}{2m}\left(|\vec{p}|^2 + m^2 \omega^2 r^2\right)$$; recall the most general quadratic case.

(c) Find $\Omega(E)$ for the linear rotator, where A is the moment of inertia.

$$\mathcal{H} = \frac{1}{2A}\left[p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} \right]$$

Note that the motion of a particle in a 1-D well, regardless of the shape of $V(x)$, at constant E, is periodic. The period is given by $\Omega(E)$. Consider:

$$\Omega(E) = \int \delta\left(E - \frac{p^2}{2m} - V(x)\right) dp\, dx$$

$$= 2 \int \frac{2m}{2} \frac{dx}{\sqrt{2m\left(E - V(x)\right)}} = 2 \int \frac{dx}{|\text{velocity at } x|} = \text{period}$$

Problem ⑤ :

Find $\Omega(E)$ for the spherical pendulum or linear molecule in an electric field:

$$\mathcal{H} = \frac{1}{2A}\left\{ p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} \right\} - \mathcal{E} d \cos\theta$$

d is the dipole moment, and, $-\mathcal{E}d < E$

The more general problem is to find $\Omega(E)$ for non-simple systems. We therefore develop some general rules. We take advantage of any separability of variables in the Hamiltonian if this exists:

$$\mathcal{H} = \underbrace{\mathcal{H}_1}_{\text{one set}} + \underbrace{\mathcal{H}_2}_{\text{another set}}$$; we assume no overlap between the two sets of variables.

Then:

$$\Omega(E) = \int \delta(E - H_1 - H_2)\, dV_1\, dV_2$$

$$= \int \Omega_2(E - H_1)\, dV_1$$; performing the integration over V_2.

We now use the trick of substituting a δ function which just multiplies by unity:

$$\Omega(E) = \int \delta(E_1 - H_1)\, \Omega_2(E - H_1)\, dE_1\, dV_1$$

$$= \int \delta(E_1 - H_1)\, \Omega_2(E - E_1)\, dE_1\, dV_1$$

or:

$$\boxed{\Omega(E) = \int \Omega_1(E_1)\, \Omega_2(E - E_1)\, dE_1}$$

This is reminiscent of the convolution integral. We have an immediate generalization to n subsystems:

$$H = \sum_{j=1}^{n} H_j$$; then:

$$\Omega(E) = \int \delta\left(E - \sum_{j=1}^{n} H_j\right) \prod_{j=1}^{n} dV_j$$

$$= \int \delta\left(E - \sum_{j=1}^{n} E_j\right) \prod_{j=1}^{n} \delta(E_j - H_j) \prod_{j=1}^{n} dV_j\, dE_j$$

Then:

$$\boxed{\Omega(E) = \int \delta\left(E - \sum_{j=1}^{n} E_j\right) \prod_{j=1}^{n} \Omega_j(E_j)\, dE_j}$$

As an example, let $n=2$:

$$\Omega(E) = \int \delta(E - E_1 - E_2)\, \Omega_1(E_1)\, \Omega_2(E_2)\, dE_1\, dE_2$$

$$= \int \Omega_1(E_1)\, \Omega_2(E - E_1)\, dE_1$$

as before. The point of the above manipulations is to get rid of phase space integrations and get to integrals over energies as quickly as possible.

Subsystem Distribution Functions:

We now consider the distribution function of some subsystem, part of a given large system. Call the Hamiltonian of the large system $\mathcal{H}$. We ask for the probability that the system lies in the volume V of the phase space. Recall the characteristic function:

$$\psi_A(q_1 \cdots p_1 \cdots) = \begin{cases} 0, & \text{pt. outside } V \\ 1, & \text{pt. inside } V \end{cases}$$

To find the requested probability, obtain the expectation value of ψ_A over the system with the microcanonical distribution function:

$$\text{Probability for system in } V = \overline{\psi_A} = \frac{1}{\Omega(E)} \int \psi_A \, \delta(E - \mathcal{H}) \, dV$$

Consider the case where ψ_A depends on the q's and p's of subsystem 1 only, assuming separability of the Hamiltonian, $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$. Then $dV = dV_1 \, dV_2$ and:

$$\overline{\psi_A} = \frac{1}{\Omega(E)} \int \psi_A \, \delta(E - \mathcal{H}_1 - \mathcal{H}_2) \, dV_1 \, dV_2$$

$$= \frac{1}{\Omega(E)} \int \psi_A \, \Omega_2(E - \mathcal{H}_1) \, dV_1$$, replacing the integration over V_2.

Suppose now that the volume for subsystem 1 is infinitesimal, that is, $V_1 \rightarrow \Delta V$. Since the characteristic function refers to subsystem 1, we may immediately write:

$$\text{Probability of subsystem 1 in } \Delta V = \frac{\Omega_2(E - \mathcal{H}_1)}{\Omega(E)} \Delta V$$

$$\equiv \rho_1 \, \Delta V$$

We see that ρ_1 is the probability density for subsystem 1.

$$P_1 = \frac{\Omega_2(E-H_1)}{\Omega(E)}$$

This relation may be surmised from the way we calculated $P(p_1)$, etc, for the perfect gas cases:

$$P(V_1) = \frac{\int \delta(E-H_1-H_2)\, dV_2}{\int \delta(E-H_1-H_2)\, dV_1\, dV_2} = \frac{\Omega_2(E-H_1)}{\Omega(E)} = P_1$$

We see we need only find the structure functions Ω's to get the probability densities P. We now ask what is the energy distribution of subsystem 1?

$$P(E_1) = \int \delta(E_1 - H_1)\, P_1\, dV_1 = \int \delta(E_1 - H_1) \frac{\Omega_2(E-H_1)}{\Omega(E)}\, dV_1$$

$$= \frac{\Omega_1(E_1)\, \Omega_2(E-E_2)}{\Omega(E)}$$

note, finally, that $P(E_1)$ is normalized:

$$\int P(E_1)\, dE_1 = \int P_1\, dV_1 = 1$$

LECTURE 10 : 10-31-61

Recall : Subsystem Functions : Structure Functions :

$$P_1 = \frac{\Omega_2(E-H_1)}{\Omega(E)}$$

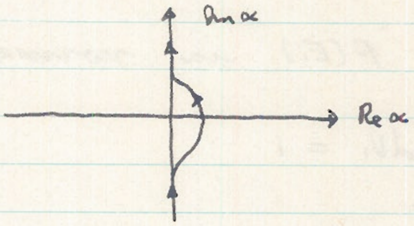
$$\Omega(E) = \int \delta(E - \Sigma E_j)\ \Pi\ \omega_j(E_j)\ dE_j$$

Previously we have used the notation $\Omega_j(E_j)$ for the subsystem structure function instead of $\omega_j(E_j)$.

Now introduce in $\Omega(E)$ our representation of the δ function:

$$\Omega(E) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} d\alpha \int e^{\alpha(E - \Sigma E_j)}\ \Pi\ \omega_j(E_j)\ dE_j$$

where we deform to the right around any poles or branch points along $\operatorname{Im}\alpha$:



We write :

$$\Omega(E) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} d\alpha\ e^{\alpha E}\ \prod_j \int e^{-\alpha E_j}\ \omega_j(E_j)\ dE_j$$

We assume the $\omega_j(E_j)$ can be bounded from above by a polynomial in any example we consider. Suppose all the subsystems are identical in form, hence :

$$\Omega(E) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} d\alpha\ e^{\alpha E}\left\{\int e^{-\alpha E}\ \omega(E)\ dE\right\}^N$$

We deform to the right because of $\int e^{-\alpha E}\omega(E)\,dE$ is defined only for $\operatorname{Re}\alpha > 0$.

We assume some sort of power behaviour for $w(E)$, so that $e^{-\alpha E}$ controls the situation for large E. Choose:

$$w(E) = B E^{l}$$

As examples, we have seen for the free particle, $l = -1/2$ in the 1-D case, $l = +1/2$ in the 3-D case. We then want to evaluate:

$$\int_0^\infty e^{-\alpha E} E^{l} dE = \frac{\Gamma(l+1)}{\alpha^{l+1}}$$

Then:

$$\Omega(E) = \frac{B^N}{2\pi i} \int_{-i\infty}^{i\infty} d\alpha \, e^{\alpha E} \left\{ \frac{\Gamma(l+1)}{\alpha^{l+1}} \right\}^N$$

$$= \frac{[B\,\Gamma(l+1)]^N}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{\alpha E}}{\alpha^{N(l+1)}} d\alpha$$

$$\boxed{\Omega(E) = [B\,\Gamma(l+1)]^N \frac{E^{N(l+1)-1}}{\Gamma(N(l+1))}}$$

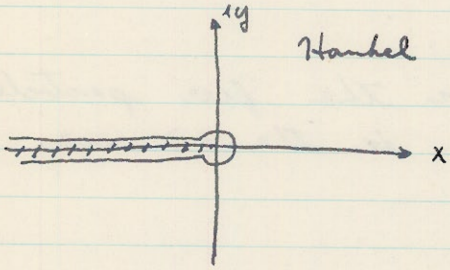
from previous results. For the case of 3-D free particle, $w(E)$ is:

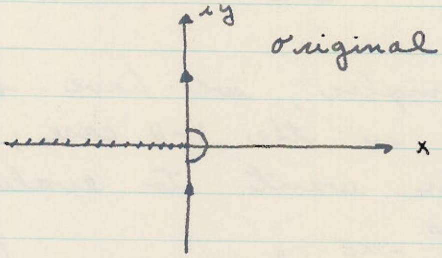
$$w(E) = 4\pi V m \sqrt{2mE}$$

One could again go to the MB limit for $\Omega(E)$ by considering asymptotic behaviour for large N. The exact answer above is not very good for applications, nor is the asymptotic one. We then consider making an approximation at the start in the hope of gaining a more tractable result, the virtue of which is to extend SM into areas where exact solutions cannot be obtained.

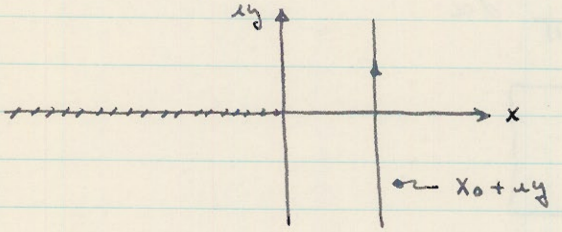
Since it is the Γ function of large argument that is giving us analytical difficulty, we will look for a way of approximating its integral definition.

Consider : $\frac{1}{2\pi i}\int_C \frac{e^z}{z^n}\,dz$, where C is either the regular or Hankel contour.





However, since the integrand is analytic everywhere except on the cut and at the origin, we can deform (do properly!) to any convenient contour. The following will be found convenient when we apply the method of steepest descents.



Now we have:

$$\frac{1}{\Gamma(n)} = \frac{1}{2\pi i}\int_{x_0 - i\infty}^{x_0 + i\infty} \frac{e^z}{z^n}\,dz \quad ; \; n \geqslant 0$$

or:

$$\frac{1}{\Gamma(n)} = \frac{e^{x_0}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{iy}}{(x_0 + iy)^n}\,dy$$

Note we have not yet made any approximations.
We see that $\frac{e^{iy}}{(x_0+iy)^n}$ has its maximum modulus at $(x_0, 0)$

That is:

$$\text{Mod. integrand} = \frac{1}{(x_0^2 + y^2)^{n/2}}$$

which is a maximum at $y = 0$. Under the proper conditions:

$$\frac{1}{(x_0^2+y^2)^{n/2}} = \frac{1}{x_0^n}\,\frac{1}{\left(1+\frac{y^2}{x_0^2}\right)^{n/2}} \simeq \frac{1}{x_0^n}\,e^{-\frac{ny^2}{2x_0^2}}$$

for $\frac{y^2}{x_0^2} \ll \sqrt{\frac{4}{n}}$, using the same reasoning as previously.

We see that the behaviour along the path about x_0 is essentially Gaussian, with almost all contributions coming from around $|y| \sim \frac{x_0}{\sqrt{n}}$. We must choose x_0 auspiciously to get the Gaussian form, however. Under these conditions, all of the integral comes from close to the real axis, suggesting that an expansion of the integrand around x_0 might prove useful. Recall:

$$\frac{1}{\Gamma(n)} = \frac{1}{2\pi i}\int_{x_0-i\infty}^{x+i\infty} \frac{e^z}{z^n}\,dz$$

Write $\frac{e^z}{z^n} = e^{f(z)}$; $f(z) = z - n\ln z$

Expanding by Taylor's theorem:

$$f(z) = f(x_0) + (z-x_0)f'(x_0) + \tfrac{1}{2}(z-x_0)^2 f''(x_0) + \cdots$$

$$= x_0 - n\log x_0 + (z-x_0)\left(1-\frac{n}{x_0}\right) + (z-x_0)^2\frac{n}{2x_0^2}$$

$$- (z-x_0)^3\frac{n}{3x_0^3} + \cdots$$

or:

$$f(z) = \underbrace{x_0 - n\log x_0}_{O(n)} + \underbrace{iy\left(1-\frac{n}{x_0}\right)}_{O(\sqrt{n})} - \underbrace{y^2\frac{n}{2x_0^2}}_{O(1)} + \underbrace{iy^3\frac{n}{3x_0^3}}_{O(\frac{1}{\sqrt{n}})} + \cdots$$

since we know that greatest contribution come from about $|y| \sim x_0/\sqrt{n}$. The term in y^4 would by this reasoning be $O(\frac{1}{n})$ and so on. These orders of magnitude are only for the range $|y| \sim \frac{x_0}{\sqrt{n}}$. We now retain terms $\geq O(1)$ and drop others since n is large, for a first approximation. The error is of $O(1/\sqrt{n})$ or about 10^{-11}. To this approximation the integral becomes:

$$\frac{1}{2\pi i}\int_{x_0-i\infty}^{x_0+i\infty} e^{f(z)}\,dz = \frac{1}{2\pi i}\int_{x_0-i\infty}^{x_0+i\infty} e^{z-n\ln z}\,dz$$

$$\frac{1}{2\pi i}\int_{x_0-i\infty}^{x_0+i\infty} e^{f(z)}\,dz = \frac{1}{2\pi i}\int_{-\infty}^{\infty} e^{x_0 - n\ln x_0 + iy\left(1-\frac{n}{x_0}\right) - \frac{ny^2}{2x_0}}$$

$$\cdot\left\{1 + i\frac{n}{3x_0^2}y^3 + \cdots\right\} i\,dy$$

$$= \frac{e^{x_0}x_0^{-n}}{2\pi}\int_{-\infty}^{\infty} e^{iy\left(1-\frac{n}{x_0}\right) - \frac{ny^2}{2x_0^2}}\,dy$$

where we have expanded the exponential in y^3, and consequently took the first term only in the spirit of the approximation. By completing the square:

$$\frac{1}{\Gamma(n)} \simeq \frac{e^{x_0}x_0^{-n}}{2\pi}\, e^{-\frac{(x_0-n)^2}{2n}}\int_{-\infty}^{\infty} e^{-\left[\sqrt{\frac{n}{2}}\frac{y}{x_0} - i\frac{x_0-n}{\sqrt{2n}}\right]^2}dy$$

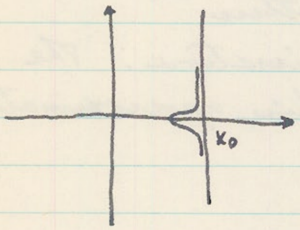
$$= \frac{e^{x_0}x_0^{-n}}{2\pi}\, e^{-\frac{(x_0-n)^2}{2n}}\; x_0\sqrt{\frac{2\pi}{n}}$$

or:

$$\frac{1}{\Gamma(n)} \simeq \frac{e^{x_0}x_0^{-(n-1)}}{\sqrt{2\pi n}}\, e^{-\frac{(x_0-n)^2}{2n}}\left\{1 + \theta\left(\frac{1}{\sqrt{n}}\right)\right\}$$

Note, however, we have in our result the saddle point parameter x_0. What we have done here is akin to approximating π and obtaining it as a function of an adjustable parameter.

What do we choose for x_0? Of course, choose it such that the greatest contributions come from close to the real axis.



Suppose we choose $x_0 = n$. Note that the integrand above does not oscillate when $x_0 = n$. Hence we can get rid of the $\theta(1/\sqrt{n})$ term as it is odd and now multiplies an even function, hence vanishing when the integral is taken. This is exactly the method of saddle points or steepest descents.

LECTURE 11 : 11-2-61

Recall:

$$\frac{1}{\Gamma(n)} \simeq \frac{e^{x_0} x_0^{-(n-1)}}{\sqrt{2\pi n}} e^{-\frac{(x_0-n)^2}{2n}}$$

Now there was a value of x_0 such that the error in $\frac{1}{\Gamma(n)}$ is of $\mathcal{O}(\frac{1}{n})$ instead of $\mathcal{O}(\frac{1}{\sqrt{n}})$. This was when $x_0 = n$, which defines the saddle point since the first derivative was seen to vanish.

$$\int \frac{e^z}{z^n} dz = \int e^{f(z)} dz \quad ; \quad f(z) = z - n\log z \; ; \; f'(z) = 1 - \frac{n}{z}$$

Thus we find it most auspicious to take the contour thru this saddle point. At $x_0 = n$, we have:

$$\frac{1}{\Gamma(n)} = \frac{e^n n^{-(n-1/2)}}{\sqrt{2\pi}} \left(1 + \mathcal{O}(\tfrac{1}{n}) + \cdots \right)$$

or:

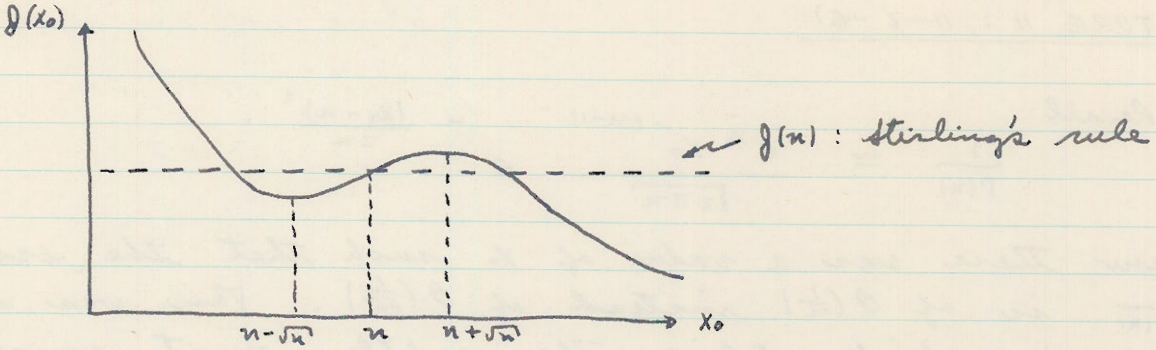
$$\boxed{\log \Gamma(n) = (n - 1/2)\log n - n + \tfrac{1}{2}\log 2\pi + \mathcal{O}(\tfrac{1}{n})}$$

This is the Stirling Approximation. However, the series does not actually converge. This is because the series is an asymptotic series where we have a minimum term in the series. What is done is to truncate the series while it still decreases.

Problem ⑥ : Find the term $\mathcal{O}(\frac{1}{n})$ in $\frac{1}{\Gamma(n)}$ when $x_0 = n$.

What happens if we miss the saddle point by a certain amount? It happens that we still get a good approximation.

set: $$g(x_0) = \frac{e^{x_0} x_0^{-(n-1)}}{\sqrt{2\pi n}} e^{-\frac{(x_0-n)^2}{2n}}$$



What happens in the region around $x_0 = n$? Take $\frac{d}{dx_0} \log f(x_0)$:

$$\frac{d}{dx_0} \log f(x_0) = 1 - \frac{n-1}{x_0} - \frac{2(x_0-n)}{2n} = (x_0-n)\left(\frac{1}{x_0} - \frac{1}{n}\right) + \frac{1}{x_0}$$

$$= -\frac{(x_0-n)^2}{x_0 n} + \frac{1}{x_0}$$

When this derivative vanishes (we look for extrema) we find two roots:

$$(x_0-n)^2 = n \quad ; \quad x_0 = n \pm \sqrt{n}$$

Professor Glauber claims that there is really very little variation around $f(n)$, the Stirling value, between $n-\sqrt{n}$ and $n+\sqrt{n}$. The extrema in the diagram above are exaggerated as $f(x_0)$ is almost flat in this region. Apparently it is not known whether or not n is an inflection point. The second logarithmic derivative goes as $-\frac{1}{n^2}$ at this point, so it is very near an inflection and becomes one in the limit of large n.

Now consider:

$$\frac{f(x_0)}{f(n)} = e^{x_0-n}\left(\frac{x_0}{n}\right)^{-(n-1)} e^{-\frac{(x_0-n)^2}{2n}}$$

and: $$\log \frac{f(x_0)}{f(n)} = x_0 - n - (n-1)\log\frac{x_0}{n} - \frac{(x_0-n)^2}{2n}$$

We want to find some conditions on the behaviour of the amount we miss the saddle point so that the higher order terms vanish as $n \to \infty$. Define the fractional change in x_0 as:

$$\Delta = \frac{x_0 - n}{n}$$

It is this quantity whose behaviour we want to determine.

$$\log \frac{f(x_0)}{f(n)} = n\Delta - (n-1)\log(1+\Delta) - \frac{n}{2}\Delta^2$$

$$= -n\left\{\log(1+\Delta) - \Delta + \tfrac{1}{2}\Delta^2\right\} + \log(1+\Delta)$$

where we will use $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$

Can we neglect $\log(1+\Delta)$? We see that $-n\log(1+\Delta)$ will be the determining factor as n grows large.

Then:

$$\log \frac{f(x_0)}{f(n)} = -n\left\{\frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \cdots\right\} + \log(1+\Delta)$$

If we retain only the Δ^3 term, then we require that

$\frac{n\Delta^3}{3} \to 0$ as $n \to \infty$ for convergence.

In mathematical notation: $|\Delta| \doteq o\left(\frac{1}{n^{1/3}}\right)$, that is,

$n^{1/3}|\Delta| \to 0$ as $n \to \infty$. Since $\Delta = \frac{x_0 - n}{n}$ and if:

$|\Delta| << \frac{1}{n^{1/3}}$, then $|x_0 - n| << n^{2/3}$

Hence the range where Stirling's Rule holds is:

$n \pm o(n^{2/3})$. Thus, if $n = 10^{21}$, the range is 10^{14} and $\Delta = 10^{-7}$.

The Partition Function:

We introduce a new quantity:

$$\boxed{\int e^{-\alpha E}\,\Omega(E)\,dE}$$

This is kind of a Laplace Transform of the structure function. Recall:

$$\Omega(E) = \int \delta\left(E - \Sigma E_j\right) \Pi\, \Omega_j(E_j)\, dE_j$$

We have sometimes used ω_j for Ω_j. Now:

$$\Omega(E) = \frac{1}{2\pi i}\int e^{\alpha E - \alpha \Sigma E_j}\, \Pi\, \Omega_j(E_j)\, dE_j\, d\alpha$$

$$= \frac{1}{2\pi i}\int e^{\alpha E}\, \Pi \int e^{-\alpha E_j}\, \Omega_j(E_j)\, dE_j\, d\alpha$$

when we introduce the representation for the Dirac δ function.

Now, as above, define the Partition Function: $\chi(\alpha)$:

$$\boxed{\chi(\alpha) = \int_0^\infty e^{-\alpha E}\,\Omega(E)\,dE}$$

We now examine some properties of $\chi(\alpha)$ using the microcanonical distribution function:

$$\chi(\alpha) = \int e^{-\alpha E}\,\delta(E - \mathcal{H})\,dV\,dE = \int e^{-\alpha \mathcal{H}}\,dV$$

We usually find the partition function easier to work with than the structure function, and more convenient.

What about muliple system structure functions? From above:

$$\chi(\alpha) = \int e^{-\alpha E}\,\Omega(E)\,dE = \int e^{-\alpha E}\,\delta\left(E - \Sigma E_j\right) \Pi\, \Omega_j(E_j)\, dE_j\, dE$$

$$\chi(\alpha) = \int e^{-\alpha \sum E_j} \prod \Omega_j(E_j)\, dE_j = \prod_j \int e^{-\alpha E_j} \Omega_j(E_j)\, dE_j$$

$$= \prod_j \chi_j(\alpha)$$

We see that for non-overlapping sub-systems, the system partition function is the product of the subsystem partition functions. This also arises from the fact that $\chi(\alpha)$ is the Laplace transform of $\Omega(E)$, which is a convolution integral, and so by a theorem of Laplace and Fourier analysis, we have $\chi(\alpha) = \prod_j \chi_j(\alpha)$. We also note that α is the same for all the subsystems, a property also had by the temperature. The introduction of Laplace transform techniques also justifies the previous use of the Dirac δ function.

Suppose we want to invert $\chi(\alpha) = \int e^{-\alpha E} \Omega(E)\, dE$ knowing $\chi(\alpha) = \prod_j \chi_j(\alpha)$. Use the Mellin theorem:

If for real α_1, $\int e^{-\alpha_1 E} |\Omega(E)|\, dE$ exists, then it is possible to invert $\chi(\alpha)$ by the following formula

$$\Omega(E) = \frac{1}{2\pi i} \int_{\alpha_2 - i\infty}^{\alpha_2 + i\infty} e^{\alpha E} \chi(\alpha)\, d\alpha$$

where $\operatorname{Re} \alpha_2 \geq \alpha_1$; $\chi(\alpha_1) = \int e^{-\alpha_1 E} |\Omega(E)|\, dE$

This is the Mellin inversion theorem. For our purposes, it is enough to take $\operatorname{Re} \alpha > 0$ or any contour in the RHP since $\chi(\alpha)$ involves nothing more pathological than polynomials. This is similar to what we have done before with the δ function, now its use is justified by the formal machinery of Laplace transforms.

We consider some more properties: Recall:

$$\chi(\alpha) = \int e^{-\alpha E} \Omega(E)\, dE = \int e^{-\alpha \mathcal{H}}\, dV$$

① $\chi(0) = \int dV = \infty$, in general.

② $\chi(\alpha) \geq 0$; α real

③ $$\frac{d\chi(\alpha)}{d\alpha} = -\int E\, e^{-\alpha E}\, \Omega(E)\, dE \leq 0$$

$\chi(\alpha)$ is analytic for $\mathrm{Re}\,\alpha > 0$, because:

$$\frac{d^n\chi}{d\alpha^n} = (-1)^n \int E^n e^{-\alpha E}\, \Omega(E)\, dE \quad \text{exists.}$$

④ $$\frac{d}{d\alpha}\log\chi(\alpha) = \frac{\chi'(\alpha)}{\chi(\alpha)} \leq 0$$; $\log\chi(\alpha)$ decreases monotonically.

⑤ $$\frac{d^2}{d\alpha^2}\log\chi(\alpha) = \frac{\chi''(\alpha)}{\chi(\alpha)} - \frac{(\chi'(\alpha))^2}{\chi^2(\alpha)} > 0$$

which can be seen from:

$$\frac{d^2}{d\alpha^2}\log\chi(\alpha) = \frac{\int E^2 e^{-\alpha E}\,\Omega(E)\,dE}{\int e^{-\alpha E}\,\Omega(E)\,dE} - \frac{\left(\int E\, e^{-\alpha E}\,\Omega(E)\,dE\right)^2}{\left(\int e^{-\alpha E}\,\Omega(E)\,dE\right)^2}$$

We now invent what we hope is a distribution function:

$$\boxed{\rho(E) = \frac{e^{-\alpha E}\,\Omega(E)}{\chi(\alpha)}} \quad ; \quad \int \rho(E)\,dE = 1$$

Then $\int E^n \rho(E)\,dE = \overline{E^n}$, if $\rho(E)$ is a distribution function. We then recognize that:

$$\frac{d^2}{d\alpha^2}\log\chi(\alpha) = \overline{E^2} - (\bar{E})^2 = \overline{(E-\bar{E})^2} = \overline{E^2} - 2\overline{E\bar{E}} + \bar{E}^2$$

and hence $\frac{d}{d\alpha}\log\chi(\alpha)$ increases monotonically.

LECTURE 12 : 11-7-61

Recall: $\Omega(E) = \int \delta(E - \sum E_j) \prod \Omega_j(E_j)\, dE_j$

$$= \frac{1}{2\pi i} \int e^{\alpha E} \prod \int e^{-\alpha E_j} \Omega_j(E_j)\, dE_j\, d\alpha$$

$$= \frac{1}{2\pi i} \int e^{\alpha E} \prod \chi_j(\alpha)\, d\alpha$$

or: $\Omega(E) = \frac{1}{2\pi i} \int e^{\alpha E} \chi(\alpha)\, d\alpha$

since: $\chi(\alpha) = \prod_j \chi_j(\alpha)$

Now we could write $\Omega(E)$ as:

$$\Omega(E) = \frac{1}{2\pi i} \int e^{f(\alpha)} d\alpha \quad ; \quad f(\alpha) \equiv \alpha E + \log \chi(\alpha)$$

What about $f(0)$? $f(0) = \infty$ because $\chi(0) = \infty$.

What about $f(\alpha)$ as $\alpha \to \infty$ along $\mathrm{Re}\,\alpha > 0$?

$$\chi(\alpha) = \int_0^\infty e^{-\alpha E} \Omega(E)\, dE \quad ; \quad \chi(\infty) \to 0 \quad ; \quad \log 0 \to -\infty ,$$

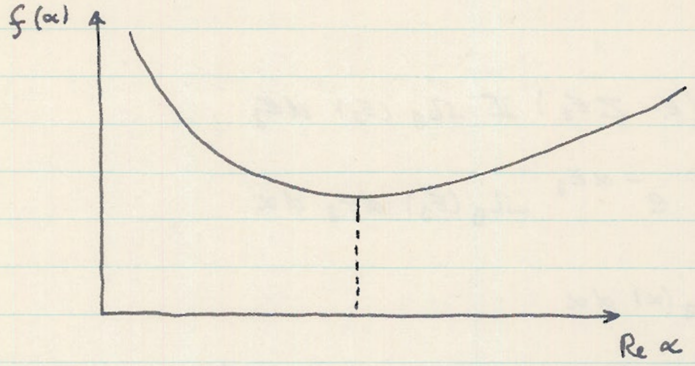
Therefore, $f(\infty) \to \infty - \infty$, and we must investigate more carefully. We note the following:

$$\chi(\alpha) > \int_0^\epsilon e^{-\alpha E} \Omega(E)\, dE > e^{-\alpha\epsilon} \int_0^\epsilon \Omega(E)\, dE = e^{-\alpha\epsilon} V(\epsilon)$$

Then: $\log \chi(\alpha) > -\alpha\epsilon + \log V(\epsilon)$
and: $f(\alpha) > \alpha(E - \epsilon) + \log V(\epsilon)$

Now choose $\epsilon < E$ and let $\alpha \to \infty$. Then along the real axis, we see that $f(\infty) = \infty$.

We now make a plot of $f(\alpha)$ versus $\mathrm{Re}\,\alpha$:



We know from the behaviour of $f(\alpha)$ at $\alpha = 0$ and ∞ that there must be at least one minimum in between. The equation for the extremum condition is:

$$E + \frac{d}{d\alpha} \log \chi(\alpha) = 0$$

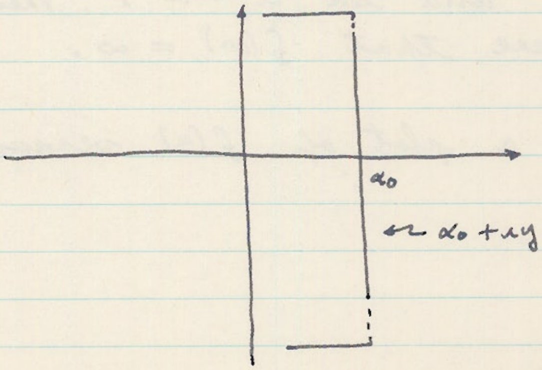
Now we know that $\frac{d}{d\alpha} \log \chi(\alpha)$ is a monotonically increasing function, hence there can be at most one root which corresponds to a minimum as can be seen from the second derivative:

$$\frac{d^2 f}{d\alpha^2} = \frac{d^2}{d\alpha^2} \log \chi(\alpha) > 0$$

Call the root of the extremum equation β; that is:

$E + \frac{\chi'(\beta)}{\chi(\beta)} = 0$, or $f'(\beta) = 0$, β being seen to give the saddle point.

We want to expand $f(\alpha)$ around β, actually a point near β called α_0 as we may not always choose to pass the contour thru the saddle point.



We assume that $\chi(\alpha)$ behaves such that contributions from deforming the contour out at $\text{Im}\,\alpha = \pm\infty$ vanishes. That is, for $\text{Re}\,\alpha > 0$, $\chi(\alpha) \to 0$ as $|\text{Im}\,\alpha| \to \infty$. Now expand $f(\alpha)$ around α_0:

$$f(\alpha) = \alpha_0 E + \log\chi(\alpha_0) + (\alpha-\alpha_0)\{\log\chi(\alpha_0)\}' + (\alpha-\alpha_0)E$$
$$+ \tfrac{1}{2}(\alpha-\alpha_0)^2\{\log\chi(\alpha_0)\}'' + \dots$$

From: $\Omega(E) = \dfrac{1}{2\pi i}\displaystyle\int_{\alpha_0-i\infty}^{\alpha_0+i\infty} e^{\alpha E}\chi(\alpha)\,d\alpha = \dfrac{1}{2\pi i}\displaystyle\int_{\alpha_0-i\infty}^{\alpha_0+i\infty} e^{f(\alpha)}\,d\alpha$,
we have:

$$\Omega(E) = \frac{e^{\alpha_0 E + \log\chi(\alpha_0)}}{2\pi}\int_{-\infty}^{\infty} e^{iy\left[E + \{\log\chi(\alpha_0)\}'\right] - \frac{1}{2}y^2\{\log\chi(\alpha_0)\}''}\,dy$$

The integral is of the form:

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}Ay^2 + iBy}\,dy = e^{-\frac{B^2}{2A}}\int_{-\infty}^{\infty} e^{-\left(\sqrt{\frac{A}{2}}\,y - \frac{iB}{\sqrt{2A}}\right)^2}\,dy = \sqrt{\frac{2\pi}{A}}\,e^{-\frac{B^2}{2A}}$$

Thus we have for $\Omega(E)$:

$$\Omega(E) = \frac{e^{\alpha_0 E}\chi(\alpha_0)}{\sqrt{2\pi\{\log\chi(\alpha_0)\}''}}\;e^{-\frac{\left[E + \{\log\chi(\alpha_0)\}'\right]^2}{2\{\log\chi(\alpha_0)\}''}}$$

What is the magnitude of the error involved in cutting off the series? Recall $\chi = \prod_j \chi_j$, $\log\chi = \sum_j \log\chi_j$. Now $\log\chi_j \sim 1$, so $\log\chi \sim N$, and it is this which motivates the use of the saddle point technique and the error will tend asymptotically to 0 as $N \to \infty$, providing α_0 is sufficiently close to the saddle point. The condition on the next term in the series must be:

$$\frac{1}{6}\left|(\alpha_0-\beta)^3\underbrace{\{\log\chi(\alpha_0)\}'''}_{O(N)}\right| \ll 1$$

or: $\dfrac{|\alpha_0-\beta|}{\beta} < \dfrac{1}{N^{1/3}}$ provides the desired asymptotic behaviour.

suppose we let $\alpha_0 \to \beta$, then:

$$\Omega(E) = \frac{e^{\beta E}\,\chi(\beta)}{\sqrt{2\pi\{\log \chi(\beta)\}''}} \quad , \text{ to order } \frac{1}{N}$$

We now go to consider the probability density for, say, subsystem 1: Recall:

$$P_1 = \frac{\Omega_2^{(N-1)}(E-H_1)}{\Omega^{(N)}(E)}$$

We can now approximate this with the methods developed above. However, we have a problem with the different saddle points for the numerator and denominator. We done the case for the denominator $\Omega(E)$, but the saddle point for the numerator will be slightly different. It will turn out that we will use the denominator saddle point for the numerator and this is the reasoning for showing that missing the saddle point slightly it will not affect the result appreciably.

We also assume all subsystems identical, each with the same partition function $\xi(\alpha)$. For the denominator:

$$E + \frac{d}{d\alpha}\log\chi = E + N\frac{d}{d\alpha}\xi(\alpha)$$

$$\text{or: } \quad E + N\frac{\xi'(\beta)}{\xi(\beta)} = 0 \quad ; \quad \frac{\xi'(\beta)}{\xi(\beta)} = -\frac{E}{N}$$

When we go to find saddle point of numerator:

$$E - H_1 + (N-1)\frac{\xi'(\beta')}{\xi(\beta')} = 0$$

or:

$$\frac{\xi'(\beta')}{\xi(\beta')} = -\frac{E-H_1}{N-1} \simeq -\frac{E}{N} + \frac{H_1}{N} - \frac{E}{N}\cdot\frac{1}{N} + \cdots$$

by long division.

now suppose $\beta' = \beta + \delta\beta$; what is $\delta\beta$?

Form: $\{\log \xi(\beta')\}' = (\log \xi(\beta))' + \delta\beta\,\{\log \xi(\beta)\}'' + \cdots$

$$= -\frac{E}{N} + \delta\beta\,\{\log \xi(\beta)\}'' = -\frac{E}{N} + \frac{H_1}{N} - \frac{E}{N^2}$$

Then:

$$\delta\beta = \frac{1}{\{\log \xi(\beta)\}''}\left\{\frac{N_1}{N} - \frac{E}{N^2}\right\}$$

$$= \frac{E}{N\{\log \xi(\beta)\}''}\left\{\frac{H_1}{E} - \frac{1}{N}\right\}$$

It is clear that $\delta\beta$ is small if the $\{\;\}$ term is small. $\dfrac{E}{N\{\log \xi(\beta)\}''}$ does not change as $N \to \infty$, since E depends on N.

In the $\{\;\}$ term, $\frac{1}{N}$ is obviously small, but what about H_1/E? We must restrict H_1 such that the energy of subsystem 1 is a small fraction of the total energy E. That is, we restrict H_1/E to go as $1/N$. This can always be realized in practice. Hence we can use the same contour for both denominator and numerator, that is, make the contour of the numerator pass thru the saddle point of the denominator. Note: if the subsystems are not identical some trouble may arise. Under the above conditions, then, we can write down directly the asymptotic form of the numerator:

$$\Omega^{(N-1)}(E-H_1) = \frac{e^{\beta(E-H_1)}\prod_{j=2}^{N}\chi_j(\beta)}{\sqrt{2\pi\sum_{j=2}^{N}\{\log\chi_j(\beta)\}''}}\; e^{-\frac{\left[(E-H_1)+\sum_{j=2}^{N}\{\log\chi_j(\beta)\}'\right]^2}{2\sum_{j=2}^{N}\{\log\chi_j(\beta)\}''}}$$

now take ratio with:

$$\Omega(E) = \frac{e^{\beta E}\prod_{j=1}^{N}\chi_j(\beta)}{\sqrt{2\pi\sum_{j=1}^{N}\{\log\chi_j(\beta)\}''}}$$

to get P_1.

$$P_1 = \sqrt{\frac{\sum_{j=1}^{N} \{\log \chi_j(\beta)\}''}{\sum_{j=2}^{N} \{\log \chi_j(\beta)\}''}} \; \frac{e^{-\beta H_1}}{\chi_1(\beta)} \; e^{-\frac{\left[(E-H_1)+\sum_{j=2}^{N}\{\log \chi_j(\beta)\}'\right]^2}{2\sum_{j=2}^{N}\{\log \chi_j(\beta)\}''}}$$

now notice: $\sqrt{\;} \sim \sqrt{\frac{N}{N-1}} \approx 1 + \frac{1}{2N} \approx 1$, because the error involved is consistent with those that we are already making. Also:

$$E = -\sum_{j=1}^{N} \{\log \chi_j(\beta)\}'$$

so in the numerator of the Gaussian term we have the cancellation of all but one term: Then:

$$\boxed{P_1 = \frac{e^{-\beta H_1}}{\chi_1(\beta)} \; e^{-\frac{\left[H_1 + \{\log \chi_1(\beta)\}'\right]^2}{2\sum_{j=2}^{N}\{\log \chi_j(\beta)\}''}}}$$

We have found the MB distribution with a Gaussian correction term. We assume for the moment that we can drop the Gaussian subject to verification. Then:

$$\boxed{P_1 = \frac{e^{-\beta H_1}}{\chi_1(\beta)}}$$

and: $\int P_1\, dV = \int \frac{e^{-\beta H_1}}{\chi_1(\beta)}\, dV = \frac{\chi_1(\beta)}{\chi_1(\beta)} = 1$

What is the energy distribution? Recall:

$$P_1(E_1) = \frac{\Omega_1(E_1)\,\Omega_2(E-E_2)}{\Omega(E)}$$

Here we can write:

$$\boxed{P_1(E_1) = \frac{\Omega_1(E_1)\, e^{-\beta E_1}}{\chi_1(\beta)}}$$

and: $\int P_1(E_1)\, dE_1 = \frac{1}{\chi_1(\beta)} \int_0^E e^{-\beta E_1}\, \Omega_1(E_1)\, dE_1 = 1$

Now we should be able to find the mean value of quantities, for example, the energy:

$$\bar{E}_1 = \int_0^{E_1} E_1 P_1(E_1)\,dE_1 = \frac{1}{\chi_1(\beta)}\left(-\frac{d}{d\beta}\right)\int_0^{E} e^{-\beta E_1}\,\Omega_1(E_1)\,dE_1$$

$$= -\frac{1}{\chi_1(\beta)}\frac{d}{d\beta}\chi_1(\beta) = -\left\{\log \chi_1(\beta)\right\}'$$

In general:

$$\boxed{\overline{E_1^n} = \int_0^{E_1} E_1^n P_1(E_1)\,dE_1 = \frac{(-1)^n}{\chi_1(\beta)}\frac{d^n}{d\beta^n}\chi_1(\beta)}$$

LECTURE 13: 11-9-61

Recall: $P_1 = \dfrac{e^{-\beta H_1}}{\chi_1(\beta)}$.

What is the variance of the energy using this density function?

$$\overline{(E_1-\bar{E}_1)^2} = \overline{E_1^2} - \bar{E}_1^2 = \frac{\chi_1''(\beta)}{\chi_1(\beta)} - \left(\frac{\chi_1'(\beta)}{\chi(\beta)}\right)^2$$

$$= \frac{d}{d\beta}\left(\frac{\chi_1'(\beta)}{\chi_1(\beta)}\right) = \frac{d^2}{d\beta^2}\log\chi_1(\beta)$$

We can now consider the neglect of the Gaussian correction. Recall that β is defined by:

$$E + \sum_j \frac{d}{d\beta}\log\chi_j(\beta) = 0$$

where $E = \sum_j \bar{E}_j$. Then we see that:

$$P_1 = \frac{e^{-\beta H_1}}{\chi_1(\beta)}\, e^{\frac{-(H_1-\bar{E}_1)^2}{2\sum_{j=2}^{N}\overline{(E_j-\bar{E}_j)^2}}}$$

The obvious condition for neglecting the exponential is:

$$|H_1 - \overline{E}_1| << \sqrt{2 \sum_{j=2}^{N} \overline{(E_j - \overline{E}_j)^2}}$$

Recall the structure function for the single free-particle subsystem in 3-D where we used the notation $\omega(E)$.

$$\omega(E) = \text{constant} \times E^{1/2}$$

Then the single particle partition function is:

$$\chi_j(\alpha) = \xi(\alpha) = \int_0^\infty e^{-\alpha E} \omega(E)\, dE = \frac{\text{constant}}{\alpha^{3/2}}$$

Now: $\chi(\alpha) = \xi^N(\alpha) = \dfrac{\text{constant}}{\alpha^{3N/2}}$

or: $\log \chi(\alpha) = -\dfrac{3N}{2} \log \alpha + \text{constant}$

The condition for β is: $E + \{\log \chi(\beta)\}' = 0$, or here:

$$E - \frac{3N}{2\beta} = 0 \quad ; \quad \beta = \frac{3N}{2E}$$

Recalling our perfect gas law, we identify:

$$\frac{1}{\beta} = \frac{2E}{3N} = kT$$

In view of this, what is the variance of the subsystem?

$$\overline{(E_j - \overline{E}_j)^2} = \frac{d}{d\beta}\left(-\frac{3}{2\beta}\right) = \frac{3}{2\beta^2} \quad ; \text{ and, also: } \overline{E}_1 = \frac{3}{2\beta}$$

Then, our criteria for neglecting the Gaussian becomes:

$$\left| H_1 - \frac{3}{2\beta} \right| << \sqrt{2(N-1)\frac{3}{2\beta^2}}$$

or:

$$\left| H_1 - \tfrac{3}{2} kT \right| << \sqrt{3N}\, kT$$

We see that this always holds for sensible cases, and hence the Gaussian correction can be completely neglected.

Problem ⑦:

(a) Find $\chi(\alpha)$ for : 1-D harmonic oscillator

(b) " " " : 3-D " "

(c) " " " : Rigid Rotator with dipole moment d in the electric field E.

Use results of previous problems, and find that χ is simpler in form than Ω.

Problem ⑧:

Relativistic Case : $\mathcal{H} = c\sqrt{p^2 + m^2c^2} - mc^2$

(a) Find $\chi(\alpha)$ for : 1-D well, extreme relativistic case, $\mathcal{H} = c|\vec{p}|$, which is for photons

(b) " " " : General 1-D relativistic case

(c) " " " : 3-D extreme relativistic case

(d) " " " : 3-D well ; general relativistic case.

Problem ⑨:

Find $\chi(\alpha)$ for the plane pendulum:

$$\mathcal{H} = \frac{p_\theta^2}{2I} + I\omega^2(1-\cos\theta)$$

Recall that Ω gives the period of the pendulum and that the motion of a pendulum is given by an elliptic integral. Thus might have to find the inversion over an elliptic function to get from Ω to χ.

We now return to a discussion of the phase space density function for a subsystem:

$$P_1 = \frac{e^{-\beta H_1}}{\chi_1(\beta)}$$

Note that we have said nothing about the size of the subsystem or the number of degrees of freedom involved (If the subsystems are molecules, we have shown that we get the MB result).

We may ask how far the results on subsystems will carry. When the subsystems are small, the results can be checked experimentally, however, very large subsystems are not subjected to experimental verification.

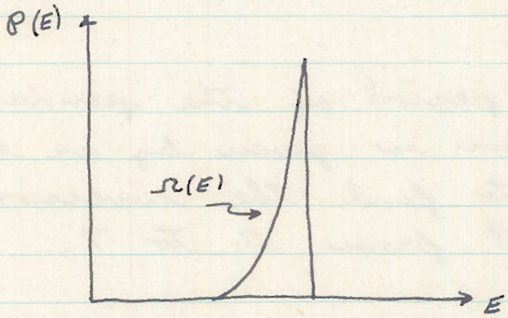
Our method of asymptotic representation has only been used since WWII as the basis of classical Statistical Mechanics. Historically, Gibbs had sensed that macroscopic results are insensitive to the distribution function used to describe the entire system. What Gibbs postulated was to try:

$$\rho = \text{constant} \times e^{-\beta H} = e^{\alpha - \beta H}$$

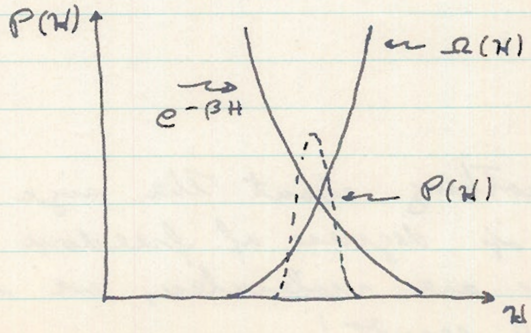
We compare this assertion with one that we have used. Consider the "solid" density function, or the density function defined in terms of the characteristic function:

$$\rho = \psi_E(\mathcal{H}) \quad ; \quad P(E) = \Omega(E)\, \psi_E(\mathcal{H}) \qquad (\text{microcanonical})$$

Recall that $\Omega(E)$ goes as $E^{\frac{n-2}{2}}$ for a free particle in an n dimensional cuble. Then the plot of $P(E)$ looks like:



Using Gibb's assertion, we have: $P(\mathcal{H}) = \text{constant} \times \Omega(\mathcal{H})\, e^{-\beta H}$



We see that if the maximum is sharp, there will be little difference between the two distributions.

As an example, take a system of N free particles. Recall that the structure function is:

$$\Omega(H) = \text{constant} \times H^{\frac{3N}{2}}$$

How does $P(\mathcal{H})$ behave? Find maximum:

$$\log P(\mathcal{H}) = \frac{3N}{2}\log \mathcal{H} - \beta H + \text{constants}$$

$$\{\log P(\mathcal{H})\}' = \frac{3N}{2E} - \beta = 0$$

Hence the maximum occurs at:

$$H = E = \frac{3N}{2\beta} = \frac{3}{2}NkT$$

We will see that this is a sharp maximum by expanding about it:

$$\log P(\mathcal{H}) = \log P(E) + \frac{1}{2}(H-E)^2\left(-\frac{3N}{2E^2}\right)$$

or: $$P(H) = P(E)\, e^{-\frac{3N}{4E^2}(H-E)^2}$$

Then, the width about $P(E)$ is approximately:

$$|H-E| \sim \frac{2}{\sqrt{3N}}E \quad ; \text{ or: } \quad \frac{|H-E|}{E} \sim \frac{2}{\sqrt{3N}}$$

Hence, Gibb's guess gives E more closely than can ever be measured. Thus, he maintained, why bother about the microcanonical distribution since the above assertion gives identical results for large systems. By using it, one should be able to infer properties of the subsystem from it. To go into this further, let us examine some properties of the exponential function and why it is so useful. The exponential of a sum is the product of the exponentials, that is:

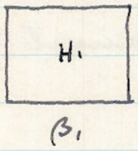
if $H = \sum_{j=1}^{N} H_j$, then $\rho = \prod_j \rho_j$; $\rho_j = \text{constant} \times e^{-\beta H_j}$

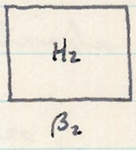
this property, resulting from the Gibbs' assumption, originally caused much controversy. Gibbs called this the method of canonical ensembles. Gibbs also felt that β was connected with temperature since the same β corresponds to all subsystems implying an equilibrium caused by constant temperature. Also the MB distribution results from the Gibbs canonical ensemble function, that is:

$$P_1 = \text{constant} \times e^{-\beta H_1} \prod_{j=2}^{N} \int e^{-\beta H_j} \, dV_j = \text{constant} \times e^{-\beta H_1}$$

This is really not as trivial as it looks, as one is tempted that the MB was used to derive the MB; however, we began not with the MB but with a distribution function <u>approximating</u> the total system. The calculations that may be done using distribution functions are obviously much simpler with the Gibbs function than with the microcanonical distribution function.

A few more words about the β to temperature relationship. Assume two perfectly isolated systems, each characterized by their own β and H.





Working on the Gibbs assertion, we would expect the total distribution function to be:

$$\rho_{12} = \text{constant} \times e^{-(\beta_1 H_1 + \beta_2 H_2)}$$

Now suppose some very slight interaction. We find that H_1, H_2 by themselves are no longer individual constants of the motion, but that now $H_1 + H_2$ is. On the other hand, if $\beta_1 \neq \beta_2$, we do not have a static distribution function, thus we must have $\beta_1 = \beta_2$ and:

$$\rho_{12} = \text{constant} \times e^{-\beta(H_1 + H_2)}$$

Hence, β must have something to do with temperature as this quantity is also characteristic of two interacting systems in equilibrium.

Gibbs gave to his distribution function the name and form:

$$\rho = e^{\alpha - \beta H} = \text{canonical ensemble}$$

It corresponds to the knowledge one has about a system at a given temperature and he surmised that $\beta = 1/kT$, knowing that using this he could derive relations analogous to the thermodynamic functions.

LECTURE 14 : 11-14-61

The way Statistical Mechanics stood around 1900 was that the only correct way to work out the principles of Statistical Mechanics was with the microcanonical distribution. However, about 1900, Gibbs made some remarkable guesses and postulated his canonical ensemble; $\rho = \text{constant} \times e^{-\beta H}$. This density function still gives a very precise ($O(\frac{1}{\sqrt{N}})$) specification of the energy and also leads to the MB subsystem result directly from the separability of the Hamiltonian, $H = \sum_j H_j$, giving $\rho_j = \text{constant} \times e^{-\beta H_j}$. The canonical ensemble was criticized on the basis that the MB seems to be assumed to get MB. The whole matter also seemed too simple to be true. However, what we really began with was a good specification of the energy in the asymptotic limit.

Gibbs also had the insight to see that β had something to do with temperature, arising from the equality of the β's for the subsystems (thermal equilibrium).

Using a relationship between β and T, Gibbs also deduced what he called functions "analogous" to the thermodynamic functions.

A great deal of mystery surrounds the canonical ensemble. It was originally considered just a calculation device (European point of view). Most Americans, naturally, feel that the foundations of Statistical Mechanics is in the canonical ensemble. What is the present attitude? We have begun with a large system, non-interacting subsystems comprising it, with specified energy and have asymptotically shown:

$$P_1 = \frac{e^{-\beta H_1}}{\chi_1(\beta)}$$

and further: $\bar{E}_1 = -\{\log \chi_1(\beta)\}'$

Thus here we have derived the canonical ensemble corresponding to part of a large system. The canonical ensemble distribution function then follows from the microcanonical distribution for a large system and then asking for the phase space density of a subsystem, obtaining the Gibbs canonical ensemble. We now go to Quantum Statistical Mechanics.

Quantum Statistical Mechanics:

Introductory Remarks: We will work mostly with the Dirac formalism. The quantum mechanical (QM) state of a system is defined by a vector is "state" or Hilbert space and is denoted by the "ket" $|j\rangle$. We can also imagine a dual space in which are defined the conjugate states denoted by the "bra" $\langle j|$. Vector scalar products are defined as: $\langle j|k\rangle$. If the set of vectors $|j\rangle$ is chosen to be orthonormal, then $\langle j|k\rangle = \delta_{jk}$.

We assume that the set $|j\rangle$ forms a complete set, that is, we are able to express any general state $|\,\rangle$ as a linear combination of the states $|j\rangle$, which now form the basis vectors of the Hilbert space. Hence:

$$|\,\rangle = \sum_j c_j |j\rangle$$

How do we find the expansion coefficients c_j? Take the scalar product of $|\rangle$ with $\langle j|$, and use the orthonormality condition:

$$c_j = \langle j | \rangle$$

We now take the scalar product of $|\rangle$ with itself and assert that the magnitudes of all state vectors can be chosen to be unity:

$$\langle | \rangle = 1 = \sum_j |c_j|^2 = \sum_j |\langle j | \rangle|^2$$

However, we note that we obtain an ambiguity in the phase of $|\rangle$. Since only closed bra-kets have physical meaning in QM, this arbitrary phase has no physical meaning, that is, $e^{i\varphi}|\rangle$ is just as good as $|\rangle$.

Interpretation of the c_j's:

The usual interpretation given to the square of the c_j's is that they give the probabilities of being in the states j, that is, we can define the probability density as:

$$\boxed{\rho_j = |c_j|^2}$$

We now mention another result of the completeness of the $|j\rangle$, that is, the outer product operator:

$$\sum_j |j\rangle\langle j| = 1$$

To show, form: $\sum_j |j\rangle\langle j|\rangle = \sum_j |j\rangle c_j = |\rangle$,

hence, $\sum_j |j\rangle\langle j| = 1$. This is also called the

resolution of unity according to the set $|j\rangle$ and also the projection operator.

We now associate with every "observable" or physical quantity an operator. Denote one of the observables by the operator A. Now in QM, we talk about eigenvalues, eigenfunctions, and eigenstates, of an operator, these being given by an "eigenvalue equation":

$$A|a'\rangle = a'|a'\rangle$$

a' is the eigenvalue of A
$|a'\rangle$ is the eigenstate of A

A similar equation holds in the dual space:

$$\langle a'|a' = \langle a'|A^{\dagger} = \langle a'|A$$

because every observable is represented by an Hermitean operator, that is, $A = A^{\dagger}$.

Now we are often concerned with the observable energy. The energy is denoted by the eigenvalues of the Hamiltonian operator $\mathcal{H}$:

$$\boxed{\mathcal{H}|E_j\rangle = E_j|E_j\rangle}$$

In solving problems in QM, it is often convenient to find sets of commuting operators. For a given problem, there are in general some largest number of commuting operators, which Dirac calls a complete set of commuting operators. In the central field problem, E_j, L^2, and L_z are the only constants of the motion and form a complete set of commuting observables. In the central field problem of classical mechanics (CM), all components of the angular momentum are constants of the motion. It is generally found for a given problem that the number of commuting observables is usually less than the classical constants of the motion. Finding the commuting observables is the most important approach in attempting to classify the states of the system.

Many times we work with time dependent states. In non-relativistic QM, time is treated as a parameter. The Schroedinger equation says that the Hamiltonian operator produces an infinitesimal translation in time:

$$\mathcal{H}|t\rangle = i\hbar \frac{\partial}{\partial t}|t\rangle$$

recalling that t is not an eigenvalue. We will usually work with a time independent Hamiltonian. Then the Schroedinger equation can be solved for a general state at some general time, $t > 0$:

$$|t\rangle = e^{-\frac{i}{\hbar}\mathcal{H}t}|0\rangle$$

We suppose that we begin with an eigenstate of the energy, that is:

$$|0\rangle = |E_j\rangle \quad , \text{ then:}$$

$$|t\rangle = e^{-\frac{i}{\hbar}Ht}|E_j\rangle = e^{-\frac{i}{\hbar}E_j t}|E_j\rangle$$

However, $e^{-\frac{i}{\hbar}E_j t}$ is just a phase factor, thus a system originally in an eigenstate will remain in an eigenstate for all time if H is independent of time. Now consider, however, a general state that depends on the time:

$$|t\rangle = \sum_j c_j(t)\,|j\rangle$$

where $c_j(t) = \langle j|t\rangle$ and with $\rho_j(t) = |c_j(t)|^2$

Following the usual interpretation of QM, we define the <u>expectation value</u> of the operator A as:

$$\langle A\rangle = \langle t|A|t\rangle = \sum_k c_k^*(t)\,\langle k|\,A\,\sum_j c_j(t)\,|j\rangle$$

$$= \sum_{jk} c_k^*(t)\,c_j(t)\,\langle k|A|j\rangle$$

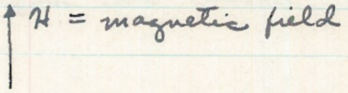
which can be interpreted as the result of repeated measurements of A.

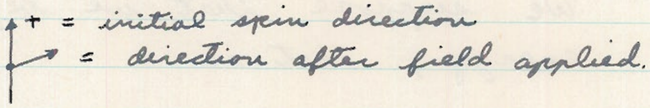
In principle, by controlling the QM system it is possible to determine the whole set $c_j(t)$ completely and reproduce it each time up to a phase factor. We say the c_j's are knowable, providing we treat the system carefully. In practice, we can rarely be this careful, so we hardly ever know the c_j's completely.

Example:

Spin states can be labelled by ± ½ values of the spin angular momentum along some direction, say the z direction.

Begin with the state described by $|c_+| = 1$, $|c_-| = 0$, or spin up. Now rotate the spin down by some resonance EM field. Then $|c_+| \neq 1$; $|c_-| \neq 0$ anymore.





This illustrates that aside from an initial phase factor, c_+ and c_- have a relative phase relation with each other that has physical meaning. However, to know what these phases are requires a great deal of information, and we would usually lose track of the phase relationship after a certain period of time.

What we are saying is that although it is possible to know the c_j's precisely and that this is consistent with QM, in reality it is difficult to prepare the c_j's as well as possible. Therefore, we introduce another type of ensemble which has to do with imperfect system preparation. Thus we must introduce the notion of an ensemble average over what already is a sort of ensemble average. This is the heart of Quantum Statistical mechanics. That is, we must in general define as the average of the observable A, the quantity:

$$\boxed{\overline{\langle A \rangle} = \sum_{jk} \overline{c_k^*(t)\, c_j(t)}\, \langle k | A | j \rangle}$$

We see that we must use two kinds of ensembles; one with the most complete QM knowledge possible, limited, of course, by the basic nature of QM, in order to get the C_j's; and another which expresses our inaccurate knowledge of initial conditions and the past history of the c_j's. Hence, in order to find the mean value of an operator, we must take the usual kind of ensemble average over the c_j's.

What this second ensemble average will result in is the notion of the density operator, analogous to a classical distribution function:

$$\boxed{\rho(t) \equiv \overline{|t\rangle\langle t|}}$$

If we take matrix elements of this operator, we find:

$$\langle k|\rho(t)|j\rangle = \overline{\langle k|t\rangle\langle t|j\rangle} = \overline{C_j^*(t)\,C_k(t)} = \overline{C_k(t)\,C_j^*(t)}$$

Then we can write:

$$\overline{\langle A\rangle} = \sum_{jk} \langle j|\rho(t)|k\rangle\langle k|A|j\rangle = \sum_{j} \langle j|\rho(t)\,A|j\rangle$$

Then:

$$\boxed{\overline{\langle A\rangle} = \text{Trace}\,\{\rho A\}}$$

LECTURE 15 : 11-16-61

Quantum Statistical Mechanics:

Recall the expansion of a state in a complete vector set:

$$| \rangle = \sum_{j} C_j | j \rangle \quad ; \quad C_j = \langle j | \rangle$$

We specify a state by the full set of C_j, $\{C_j\}$, and everything about them except the multiplicative phase factor which has no meaning, physically. Very seldom do we know all the details to determine the full set $\{C_j\}$. Knowing all the C_j's exactly is like knowing all the coordinates of the point in phase space classically. Because we do not know all about the C_j's, there is a second indeterminism in QM besides the one pointed out by Heisenberg. This second is statistical in nature.

Recall the eigenvalue problem:

$$A|a'\rangle = a'|a'\rangle$$

This is the only case in QM where we have complete determinacy. This is not true for any other operator not diagonal in $|a'\rangle$. That is, suppose the existence of some other operator B, form:

$$|a'\rangle = \sum_{b'} |b'\rangle \langle b'|a'\rangle$$

Then B on $|a'\rangle$ gives b' with probability $\langle b'|a'\rangle$.
Thus we can write in general for the expectation value of any operator A:

$$\langle A \rangle = \langle |A| \rangle = \sum_{jj'} C_j C_{j'}^* \langle j'|A|j \rangle$$

However, when we can't express the C_j's exactly, that is, we do not know the complete past history, we must take the statistical average:

$$\overline{\langle A \rangle} = \sum_{jj'} \overline{C_j C_{j'}^*} \langle j'|A|j \rangle = \mathrm{Trace}\,(\rho A)$$

where we define a quantity known as the density operator:

$$\rho(t) = \overline{|t\rangle\langle t|} \quad ; \quad \langle \gamma|\rho|\gamma'\rangle = \overline{c_\gamma c^*_{\gamma'}}$$

Consider the case where $A = 1$:

$$\langle 1\rangle = 1 = \text{Trace}\,\rho = \sum_\gamma \langle\gamma|\rho|\gamma\rangle = \sum_\gamma \overline{|c_\gamma|^2}$$

so that the diagonal elements of ρ must have an interpretation as probabilities, that is, the probability that the system is in the state γ is given by:

$$P_\gamma = \overline{|c_\gamma|^2} = \langle\gamma|\rho|\gamma\rangle$$

<u>Example of Application</u>: System of Spin ½:

Recall the Stern - Gerlach experimental system in which an inhomogeneous magnetic field seperates spinning charged particles into two separate beams, one with spin parallel to the field and the other antiparallel. If we pass one of these beams thru another Stern - Gerlach system with the field oriented parallel to the spins of the incident beam, 100% transmission will occur. If the field is antiparallel, we have 0% transmission. This experience is analogous to the polarizer - analyzer systems used in optics.

We now assume we have prepared a system with all the spins oriented in the same direction. We will define a set of coordinates with respect to this beam. We first recall some properties of the spin operators:

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \quad ; \quad \vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z) \quad ; \quad \sigma_x\sigma_y = -\sigma_y\sigma_x \text{ , etc.}$$

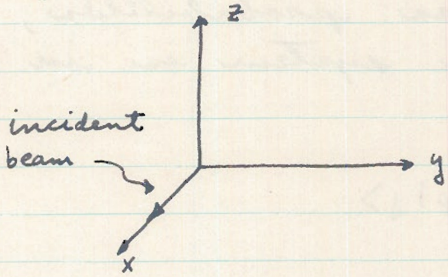
$$(i\sigma_y)^2 = -1 \quad ; \quad \sigma_x\sigma_y = i\sigma_z \text{ , etc.}$$

We take as our complete set of eigenstates, those states parallel and antiparallel to the z axis, that is, we quantize the system with respect to the z axis such that:

$$\sigma_z |+\rangle = |+\rangle \quad ; \quad \sigma_z |-\rangle = -|-\rangle$$

We see that $|+\rangle$ represents the parallel state and $|-\rangle$ the antiparallel.

However, now consider all the spins in the incident beam to be along +x direction:



How can we choose a new representation, say, $|+\rangle'$, so that σ_x instead of σ_z is diagonal in the new representation. Try forming:

$$|+\rangle' = e^{-i\pi/4\,\sigma_y} |+\rangle$$

and applying σ_x. Keep in mind $'\langle +|+\rangle' = 1$. $e^{-i\pi/4\,\sigma_y}$ is the spin rotation operator for a rotation of 90° about the y axis. We see that this operation brings z into coincidence with x. Use Euler's relation:

$$e^{i\theta\sigma_y} = \cos\sigma_y\theta + i\sin\theta\sigma_y$$

$$\cos\sigma_y\theta = 1 - \frac{(\sigma_y\theta)^2}{2!} + \frac{(\sigma_y\theta)^4}{4!} - \cdots = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots$$

$$= \cos\theta$$

$$\sin\sigma_y\theta = \sigma_y\theta - \frac{(\sigma_y\theta)^3}{3!} + \frac{(\sigma_y\theta)^5}{5!} - \cdots = \sigma_y\left[\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right]$$

$$= \sigma_y \sin\theta$$

Then:

$$\sigma_x |+\rangle' = \sigma_x \, e^{-i \pi/4 \, \sigma_y} |+\rangle$$

$$= \sigma_x \left[\cos \pi/4 - i \sigma_y \sin \pi/4 \right] |+\rangle$$

$$= \sigma_x \left[\frac{\sqrt{2}}{2} - i \sigma_y \frac{\sqrt{2}}{2} \right] |+\rangle$$

$$= \left\{ \sigma_x \frac{\sqrt{2}}{2} + \sigma_z \frac{\sqrt{2}}{2} \right\} |+\rangle = \left\{ -i \sigma_y \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \right\} \sigma_z |+\rangle$$

$$= e^{-i \pi/4 \, \sigma_y} |+\rangle$$

or: $\sigma_x |+\rangle' = |+\rangle'$, so σ_x is indeed diagonal in the $|+\rangle'$ representation. In general, we can find the eigenstates of the other spin operators by rotation.

What we want to do now is find the resolution of $|+\rangle'$ along $|+\rangle$:

$$|+\rangle' = \left(\cos \pi/4 - i \sigma_y \sin \pi/4 \right) |+\rangle$$

$$= \frac{\sqrt{2}}{2} |+\rangle - i \frac{\sqrt{2}}{2} \sigma_y |+\rangle$$

Now recall: $A|j\rangle = \sum_k |k\rangle\langle k|A|j\rangle$, or, applying to the case at hand:

$$\sigma_y |+\rangle = \sum_{j=\pm} |j\rangle\langle j|\sigma_y|+\rangle = |+\rangle\langle +|\sigma_y|+\rangle + |-\rangle\langle -|\sigma_y|+\rangle$$

Now introduce the well-known representation for the σ's with σ_z diagonal:

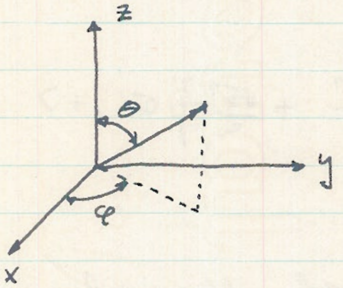
$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad ; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad ; \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

We immediately see: $\langle +|\sigma_y|+\rangle = 0$; $\langle -|\sigma_y|+\rangle = i$, so:

$$\sigma_y |+\rangle = i |-\rangle, \text{ and:}$$

$$|+\rangle' = \frac{\sqrt{2}}{2} |+\rangle + \frac{\sqrt{2}}{2} |-\rangle$$

which we immediately see corresponds to polarization in the $+x$ direction. The c_j's are the $\frac{\sqrt{2}}{2}$ and we know them exactly because we know the initial preparation of the system. In general, if we had prepared the system so that the spins were polarized in some θ, φ direction, the eigenstate of the system would be given by:



$$|+\rangle' = e^{i\varphi/2} \cos\frac{\theta}{2} |+\rangle + e^{-i\varphi/2} \sin\frac{\theta}{2} |-\rangle$$

Now suppose we are measuring some function of the spin variable: $F = F(\sigma_z)$, choosing the function to depend on the spin variable or operator σ_z for the moment. Then:

$$\langle F(\sigma_z)\rangle = \langle\, | F(\sigma_z) |\, \rangle = \left[\langle +| C_+^* + \langle -| C_-^*\right] F(\sigma_z) \left[C_+ |+\rangle + C_- |-\rangle\right]$$

$$= |C_+|^2 F(1) + |C_-|^2 F(-1)$$

Now, more generally, F is not simply a function of σ_z, but rather of $\vec{\sigma}$ as in radiation problems:

$$F = F(\sigma_x, \sigma_y, \sigma_z) = F(\vec{\sigma})$$

Then:

$$\langle F(\vec{\sigma})\rangle = \langle\, | F(\vec{\sigma}) |\, \rangle = |C_+|^2 \langle +|F|+\rangle + |C_-|^2 \langle -|F|-\rangle$$

$$+ C_+^* C_- \langle +|F|-\rangle + C_-^* C_+ \langle -|F|+\rangle$$

Now if the system is fully prepared as the result of a Stern-Gerlach experiment for example, we then can find C_-, C_+ exactly as we have shown.

However, now consider a system of unpolarized electrons, of which we assume 1/2 oriented along one direction and 1/2 opposite. We can always choose this direction parallel to the z axis or rotate the coordinate system so this is true.

The natural mean value of $F(\vec{\sigma})$ is then:

$$\tfrac{1}{2}\langle +|F|+\rangle + \tfrac{1}{2}\langle -|F|-\rangle$$

What we have done is to choose two directions arbitrarily. This is then a <u>mixture</u> of <u>pure</u> states. The result above represents $\overline{\langle F\rangle}$ in which $\overline{|C_+|^2} = \frac{1}{2}$, $\overline{|C_-|^2} = 1/2$, $\overline{C_+^* C_-} = \overline{C_-^* C_+} = 0$, providing $\pm z$ are the arbitrary directions.

Now suppose we define the operator:

$$\rho = \tfrac{1}{2}\left[|+\rangle\langle +| + |-\rangle\langle -|\right]$$

This must be the density operator because we see:

$$\overline{\langle F\rangle} = \text{Trace}\,(\rho F) = \langle +|\rho F|+\rangle + \langle -|\rho F|-\rangle$$

$$= \tfrac{1}{2}\langle +|F|+\rangle + \tfrac{1}{2}\langle -|F|-\rangle$$

Now consider the case of partial polarization of the incoming beam of charged particles. In analogy with the above case, we take for the density operator:

$$\rho = \alpha|+\rangle\langle +| + \beta|-\rangle\langle -| = \tfrac{1}{2}(1+\lambda)|+\rangle\langle +| + \tfrac{1}{2}(1-\lambda)|-\rangle\langle -|,$$

$\alpha, \beta > 0$; $\alpha + \beta = 1$, and where λ is the % polarization in the $+z$ direction.

In general, a pure state is absolutely different from a mixture. The difference usually arises in the nature of the preparation of a given experiment; that is, whether or not the experiment is completely prepared or is independent of past history and random.

We now prove the invariance of the trace under the unitary transformation U; $U^+ = U^{-1}$. We have:

$$|j\rangle' = U|j\rangle = \sum_k |k\rangle U_{kj}$$

Form: $\langle j|A|k\rangle' = \langle j|U^{+}AU|k\rangle$

We now take the trace in the new system:

$$\text{Trace}\,(\rho A)' = \text{Trace}\left\{U^{+}\rho\,\underbrace{UU^{+}}_{1}AU\right\} = \text{Trace}\,(\rho A)$$

Hence, the trace has remained invariant. Of course, the difficulty of computing the trace in one system may differ vastly from computing the trace in a representation in which both ρ and A are diagonal.

Time Dependence of the Density Operator:

We now return to a discussion of the density operator as a function of time. Recall:

$$\rho(t) = \overline{|t\rangle\langle t|}$$

What happens when we apply $i\hbar\frac{\partial}{\partial t}$ to $\rho(t)$?

$$i\hbar\frac{\partial}{\partial t}\rho(t) = \overline{\left(i\hbar\frac{\partial}{\partial t}|t\rangle\right)\langle t|} - \overline{|t\rangle\left(-i\hbar\frac{\partial}{\partial t}\langle t|\right)}$$

$$= H\,\overline{|t\rangle\langle t|} - \overline{|t\rangle\langle t|}\,H = H\rho - \rho H = [H,\rho]$$

recalling the Schroedinger equation, $i\hbar\frac{\partial}{\partial t}|t\rangle = H|t\rangle$. We note the similarity to the Liouville equation:

$$\frac{\partial\rho}{\partial t} = -\{\rho, H\} \quad (\text{classical})$$

$$\frac{\partial\rho}{\partial t} = -\frac{1}{i\hbar}[\rho, H] \quad (\text{quantum})$$

hence in the classical limit, $\frac{1}{i\hbar}[\,,\,] \to \{\,,\,\}$. Consider now the Hiesenberg picture in which operators $A(t)$ obey:

$$i\hbar\frac{dA(t)}{dt} = [A,H] + i\hbar\frac{\partial A}{\partial t}$$, called the Hiesenberg equation of motion.

This shows that even in QM, $\frac{d\rho(t)}{dt} = 0$.

LECTURE 16: 11-21-61

The density operator of QM replaces the density function of Classical mechanics:

$$\rho(t) = \overline{|t\rangle\langle t|}$$

$$i\hbar \frac{\partial \rho}{\partial t} = -[\rho, H]$$

Recall that classically we chose the microcanonical distribution for ρ and showed that averaging processes were not particularly dependent on this choice. We will do much the same thing in QM.

We have for the equation of motion of an operator in the Hiesenberg picture:

$$\frac{dA}{dt} = \underbrace{\frac{1}{i\hbar}[A, H]}_{\substack{\text{dynamical} \\ \text{part: } q, p \\ \text{changing with} \\ \text{time}}} + \underbrace{\frac{\partial A}{\partial t}}_{\substack{\text{explicit} \\ \text{time} \\ \text{dependence}}}$$

We can apply this equation to the density operator $\rho = \rho(q(t) \cdots p(t) \cdots)$, and get the Liouville equation. We use the fact that $d\rho/dt = 0$. Hence:

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{1}{i\hbar}[\rho, H] = 0}$$

We will deal with static density operators, that is, $\frac{\partial \rho}{\partial t} = 0$, or, ρ must commute with H:

$$[\rho, H] = 0$$

This will be the jumping off point for choosing ρ.

We now ask what else commutes with H? In the case of the central field, $\vec{L}^2$, L_z commute with H and form a complete set of commuting observables.

Hence ρ can be any function of the complete set of commuting observables including H. As in classical mechanics, we cannot make statements about other constants of the motion. Here as in classical theory we will not say anything about other members of the set, thus we will choose ρ as a function of H.

$$\boxed{\rho = \rho(H)}$$

This is done because the Hamiltonian characterizes all systems while the others like $\vec{L}^2$ do not always. We write, using the unit operator:

$$\rho(H) \cdot 1 = \rho(H) \sum_{E_j, k} |E_j k\rangle \langle E_j k|$$

E_j = eigenvalue of a discreet set
k = degeneracy index

We shall always deal with a discreet spectrum.
Now:

$$\rho(H) = \sum_{E_j k} |E_j k\rangle \rho(E_j) \langle E_j k|$$

and: $\langle E_j k | \rho | E_j', k' \rangle = \delta_{E_j, E_j'} \; \delta_{kk'} \; \rho(E_j)$

Hence ρ and H are diagonal in the same representation as they are commuting observables. Note that there are no matrix elements connecting existing degenerate states.

For averaging purposes, recall that we require the normalization condition:

$$\text{Trace}\, \rho = 1$$

or: $\text{Trace}\, \rho = \sum_{E_j, k} \langle E_j k | \rho | E_j k \rangle = \sum_{E_j k} \rho(E_j) = 1$

k now just gives the multiplicity of each E_j, hence each degenerate state is given equal weight.

Now recall:

$$\langle \gamma | \rho | \gamma' \rangle = \overline{C_\gamma C_{\gamma'}^*}$$

$$\text{or:} \quad \langle E_\gamma \kappa | \rho | E_{\gamma'} \kappa' \rangle = \overline{C_{E_\gamma \kappa'} C^*_{E_{\gamma'} \kappa}} = \delta_{E_\gamma E_{\gamma'}} \delta_{\kappa\kappa'} \rho(E_\gamma)$$

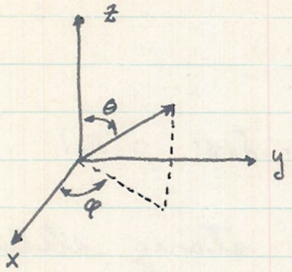
By asserting $\rho = \rho(H)$, we have made a strong statement about $\overline{CC_i^*}$, that is, they vanish when the representation used to express ρ is the eigenstate representation of the Hamiltonian because only the diagonal elements will remain. To explain this vanishing of the off-diagonal elements, write:

$$\overline{C_{E_\gamma \kappa} C^*_{E_{\gamma'}\kappa'}} = \overline{|C_{E_\gamma \kappa}| |C_{E_{\gamma'}\kappa'}| e^{i(\varphi_{\gamma\kappa} - \varphi_{\gamma'\kappa'})}}$$

Now, if the past history and initial conditions are forgotten, we lose track of the phase factors and they become uniformly distributed, hence the average over them vanishes. Statistically we say that we have no knowledge of the phases. This is precisely what $[\rho, H] = 0$ says.

We have done essentially the same procedure here as in classical mechanics when we say that we will give equal weight to degenerate states. State counting in QM corresponds to phase space volume determination classically. In classical mechanics we talked about 2F-1 constants of the motion, while in QM we talk about complete sets of commuting observables. In the central force problem, we have 3 not 5 QM constants of the motion as the various components of the angular momentum do not commute among themselves. Classically, we said that other constants of the motion than H were randomly distributed and this is also assumed in QM. This is all a consequence of commuting with H.

Consider the spin model again. Take z direction quantized, denoted by $|+\rangle$ and $|-\rangle$.



$$|\rangle = e^{i\frac{\varphi}{2}} \cos\frac{\theta}{2} |+\rangle + e^{-i\frac{\varphi}{2}} \sin\frac{\theta}{2} |-\rangle$$

We have generally for a randomly prepared, partially polarized system:

$$\rho = \alpha |+\rangle\langle+| + \beta |-\rangle\langle-| \quad ; \quad \alpha + \beta = 1$$

If there is no magnetic field present, the + and − states are degenerate and hence for a completely randomly prepared system we must deal with an unpolarized beam, or $\alpha = \beta = 1/2$. We take a field in the z direction and the Hamiltonian to be diagonal in this representation since:

$$H = -\mu \vec{\sigma} \cdot \vec{\mathcal{H}} \quad ; \quad \vec{\mathcal{H}} = (0, 0, \mathcal{H}_z)$$

since $\rho(H) = \sum_{E_j k} \rho(E_j) |E_j k\rangle\langle E_j k|$, here we have:

$$\rho(-\mu \sigma_z \mathcal{H}_z) = \rho(-\mu \mathcal{H}_z) |+\rangle\langle+| + \rho(\mu \mathcal{H}_z) |-\rangle\langle-|$$

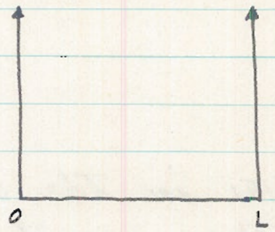
so that: $\alpha = \rho(-\mu \mathcal{H}_z)$; $\beta = \rho(\mu \mathcal{H}_z)$; μ = magnetic moment

We now display the correspondence between operations in classical statistical mechanics and quantum statistical mechanics:

Classical		Quantum
Distribution function	⟷	Density operator
Integration over Phase Space	⟷	Sum over states
Energy Integrals	⟷	Sum over energies of discreet states

For the quantum mechanical system in the laboratory, the energy levels may become extremely closely spaced or quasi-continuous.

Consider first the problem of a particle in a box.



The eigenstates in the coordinate representation are given by:

$$e^{i\vec{k}\cdot\vec{r}} \; ; \; \vec{k} = \frac{2\pi}{L}(n_1, n_2, n_3)$$

$$E = \frac{\hbar^2}{2m} \frac{(2\pi)^2}{L^2} \left(n_1^2 + n_2^2 + n_3^2 \right)$$

Suppose we put in more particles and begin to fill up to the higher states. We see a tremendous amount of degeneracy develops because different combinations of n_1, n_2, n_3 give the same sum of squares. What is an order of magnitude for the fineness of the level structure in a 1 cm³ box of protons? Take $\Delta |n|^2 \sim 1$:

$$\Delta E = \text{level spacing} \sim \underbrace{\left(\frac{m_e}{M_p}\right)}_{10^{-4}} \underbrace{\frac{\hbar^2 (2\pi)^2}{2 m_e a_B^2}}_{10\text{ ev}} \underbrace{\left(\frac{a_B}{L}\right)^2}_{10^{-16}} \sim 10^{-19}\text{ ev}$$

What we now want to show is that adding 1ev to a many particle gas is just lost in the muck. First, some conversions:

$T = 300°\text{ K} : \; kT = 1/40 \text{ ev} = .025 \text{ ev}$

$T = 1°\text{ K} : \; kT \approx 10^{-4} \text{ ev}$

kT is the average energy per particle ($U = 3NkT$). Adding 1ev to 10^{19} atoms/cc raises the energy per atom 10^{-19} ev or 10^{-15} degrees. The above digressions indicate the effect of the container size on the level density of the system.

We now proceed directly to find the QM analog of the microcanonical distribution function:

$$\text{CM}: \; \rho_E = \text{constant} \times \delta(E - \mathcal{H}) = \frac{\delta(E-H)}{\Omega(E)}$$

where $\Omega(E)$ is the structure function.

In QM, we choose by analogy, remembering the discreet nature of the energy states:

$$\boxed{\rho_E = \rho_E(H) = \frac{\delta_{E,H}}{W(E)}}$$

where $\delta_{E,H}$ is the Kronecker δ and $W(E)$ is the QM analogy to the structure function. Apply this operator to the unit operator:

$$\frac{\delta_{E,H}}{W(E)} = \frac{\delta_{E,H}}{W(E)} \sum_{E'k'} |E'k'\rangle\langle E'k'|$$

We define the Kronecker δ operator in the way it operates on the states:

$$\frac{\delta_{E,H}}{W(E)} = \frac{1}{W(E)} \sum_{E'k'} |E'k'\rangle\, \delta_{E,E'}\, \langle E'k'|$$

or: $$\rho_E(H) = \frac{1}{W(E)} \sum_{k} |Ek\rangle\langle Ek|$$

Using: $\overline{\langle A\rangle} = \text{Trace}\,\rho A$ and taking $A = 1$, we get:

$$1 = \frac{1}{W(E)}\,\text{Trace} \sum_k |Ek\rangle\langle Ek| = \frac{\text{degree of degeneracy}}{W(E)}$$

Therefore, $W(E)$ is the degree of degeneracy and is an integer. Although this is the analog to the classical microcanonical distribution function, we will not use it in this form. It is not usually $W(E)$ that we want because the degree of degeneracy, among other things, does not behave with the energy in a regular fashion and is generally not very useful at all. Also degeneracies can be completely removed, sometimes, by even the slightest perturbation.

LECTURE 17 : 11-27-61

Recall the choice for the density operator based on the analogy to the classical microcanonical distribution function.

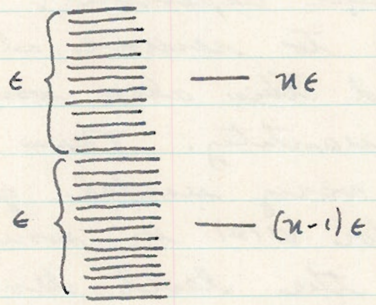
$$\rho_E = \frac{\delta_{E,H}}{W(E)} = \frac{1}{W(E)} \sum_k |Ek\rangle\langle Ek|$$

W(E) gives the degeneracy of E. However, because of the general uselessness of the degeneracy, perhaps we should not try to preserve the analogy to the microcanonical distribution function. For example, if we fill a cube in momentum space, we do not obtain any degeneracy, and also, in situations where degeneracy does exist, it does not follow any simple rule from state to state. The presence of degeneracy depends on the shape of the container to a large degree and we would not want to define a basic relation of SM on quantities that are shape dependent. Any small asymmetries also tend to remove all the degeneracies or much of it and this also would be inconvenient in any basic quantity. Thus we conclude that W(E) is not a very useful quantity. What we would really like for W(E) is something like the density of states or the level density. Precise specification of the states is not needed nor desired because of the uncertainty principle which relates quantum numbers and phase. Recall that in order to describe a classical oscillator, one supperposes states of all energies of the QM oscillator in order to get the familiar behaviour. In the transition to the classical limit, an individual quantum state has no weight since we have superposed so many different energy states.

We would want any analog of the microcanonical distribution function in QM to have the CM microcanonical distribution function (MCDF) as the classical limit. However, when we pursue this too zealously, W(E) ends up with little useful meaning.

What we must say is that there is some intrinsic uncertainty in the energy. A great many levels are included in the transition to the classical limit. There is nothing in CM that says we can specify the energy precisely. In going to the limit, we take a great many levels within the width of the uncertainty principle and then let $\hbar \to 0$, or vice versa.

How then do we formulate a useful QM treatment? We will deal with energy intervals rather than the levels themselves. Recall that the addition of energy to a system of many particles becomes spread out over all states. Then define an energy interval ϵ, large enough to include many energy levels ($\approx 10^{19}$), so that we have enough to permit the application of statistical methods. We define an indefiniteness in the energy so that for $(n-\frac{1}{2})\epsilon < E < (n+\frac{1}{2})\epsilon$ we will say that E is $n\epsilon$. Let $W_\epsilon(E)$ denote the total "degeneracy" or number of levels of $E = n\epsilon$:



We assume the system has equally spaced energy levels.

We now write:

$$\rho_\epsilon = \frac{1}{W_\epsilon(E)} \sum_k |E k\rangle\langle E k|$$

k now numbers all states in the interval $(n-\frac{1}{2})\epsilon < E < (n+\frac{1}{2})\epsilon$ as if E were degenerate. We still stipulate that $\text{Trace}\ \rho_\epsilon = 1$, or:

$$W_\epsilon(E) = \text{number of states in interval } \epsilon,$$

so $W_\epsilon(E)$ is no longer the degeneracy. This quantity, $W_\epsilon(E)$, if large enough, should vary smoothly from one interval to another in the asymptotic limit. When one sees how this becomes possible, we see that we can get the structure function. We do not consider any examples at this time.

We would expect the level or state density to be given by $\underline{W_\epsilon(E)}$ which will turn out to be the analog of the $^\epsilon$ structure function. If we have a set of independent subsystems, labelled j, all each in the state E_j, the total multiplicity goes as $\prod_j W_j(E_j)$, dropping the ϵ for the moment. However, the total multiplicity must conform to the constraint $E = \sum_j E_j$. Thus, including this constraint, the total multiplicity can be written as:

$$\boxed{W(E) = \sum_{\{E_j\}} \delta_{E,\, \sum_j E_j} \prod_j W_j(E_j)}$$

Thus we see that $W(E)$ is precisely analogous to the classical structure function.

Now, following the CM example, we claim that the probability that subsystem #1 occupies a particular state corresponding to energy E_1 is, as before, a ratio of two structure functions:

$$P^{(1)}_{E_1} = \frac{W^{(N-1)}(E-E_1)}{W^{(N)}(E)}$$

Now, by decomposition of the structure function, we have:

$$W^{(N)}(E) = \sum_{E_1} W^{(1)}(E_1)\, W^{(N-1)}(E-E_1)$$

Divide by $W^{(N)}(E)$:

$$1 = \sum_{E_1} \frac{W^{(1)}(E_1)\, W^{(N-1)}(E-E_1)}{W^{(N)}(E)} = \sum_{E_1} W^{(1)}(E_1)\, P^{(1)}_{E_1}$$

and we see that $\sum_{E_1} W^{(1)}(E_1)\, P^{(1)}_{E_1}$ is a probability distribution function. We can then write a distribution function for the energy:

$$P_1(E_1) = W^{(1)}(E_1)\, P^{(1)}_{E_1} = W^{(1)}(E_1)\, \frac{W^{(N-1)}(E-E_1)}{W^{(N)}(E)}$$

We have shown that W corresponds to Ω. It is now an obvious extension to define an analogy to the partition function X. Call the QM partition function Q.

$$\boxed{Q_j(\alpha) = \sum_{E_j} e^{-\alpha E_j} W_j(E_j)}$$

This is for subsystem j. Generally:

$$Q(\alpha) = \sum_{E} e^{-\alpha E} W(E) = \sum_{E} e^{-\alpha E} \sum_{\{E_j\}} \delta_{E, \sum_j E_j} \prod_j W_j(E_j)$$

$$= \sum_{\{E_j\}} e^{-\alpha \sum_j E_j} \prod_j W_j(E_j) = \sum_{\{E_j\}} \prod_j e^{-\alpha E_j} W_j(E_j)$$

$$= \prod_j \sum_{E_j} e^{-\alpha E_j} W_j(E_j) = \prod_j Q_j(\alpha)$$

reversing Π and Σ in the usual way. Hence the total partition function of independent subsystems is the product of the individual subsystem partition functions.

We now would like to concoct a representation for the Kronecker δ function as we did for the Dirac δ function. Remember that we are dealing with levels that are multiples of each other. We choose:

$$\boxed{\delta_{E,E'} = \frac{1}{2\pi i} \oint \frac{dz}{z} \, z^{1/\epsilon (E-E')}}$$

where the integration is carried out around the unit circle. We see that the above choice satisfies the requirements of $\delta_{E,E'}$ providing E, E' are multiples of ϵ. We now change variables to get into a form nearer that of the Dirac δ function.

Let: $z = e^{\alpha \epsilon}$; $dz = \epsilon z \, d\alpha$, then:

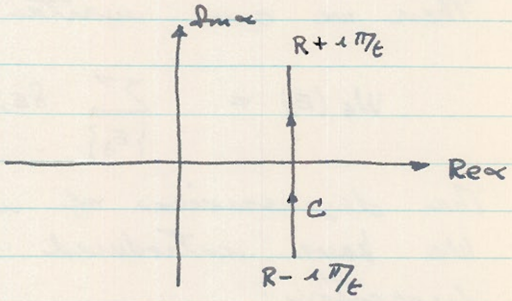
$$\delta_{E,E'} = \frac{\epsilon}{2\pi i} \int_c e^{\alpha (E-E')} \, d\alpha$$

What is the new contour?

$$\alpha = \frac{1}{\epsilon}\log z = \frac{1}{\epsilon}\log r e^{i\theta} = \frac{1}{\epsilon}\log r + i\frac{\theta}{\epsilon}$$

The old contour in the z plane around the unit circle went from, say, $-\pi$ to π, so $\operatorname{Im}\alpha$ goes from $-i\pi/\epsilon$ to $i\pi/\epsilon$. Now, although the radius of the circle in the z plane was unity, clearly it could have been any value, say, R. Then:

$$\boxed{\delta_{E,E'} = \frac{\epsilon}{2\pi i}\int_{R-i\pi/\epsilon}^{R+i\pi/\epsilon} e^{\alpha(E-E')}\,d\alpha}$$



We can now write the decomposition law for the structure function W(E) in the form:

$$W(E) = \sum_{\{E_j\}} \delta_{E,\sum_j E_j} \prod_j W_j(E_j) = \frac{\epsilon}{2\pi i}\int_C \sum_{\{E_j\}} e^{\alpha(E-\sum_j E_j)} \prod_j W_j(E_j)\,d\alpha$$

$$= \frac{\epsilon}{2\pi i}\int_C e^{\alpha E} \sum_{\{E_j\}} \prod_j e^{-\alpha E_j}\, W_j(E_j)\,d\alpha$$

$$= \frac{\epsilon}{2\pi i}\int_C e^{\alpha E} \prod_j \sum_{E_j} e^{-\alpha E_j}\, W_j(E_j)\,d\alpha$$

$$= \frac{\epsilon}{2\pi i}\int_C e^{\alpha E} \prod_j Q_j(\alpha)\,d\alpha$$

or:

$$W(E) = \frac{\epsilon}{2\pi i}\int_C e^{\alpha E}\, Q(\alpha)\,d\alpha$$

We differ from the classical case because of the finite limits and the ϵ out in front. Recall that $\frac{W(E)}{\epsilon}$ is the state density.

LECTURE 18 : 11-30-61

We have constructed the quantum-mechanical analog to the classical MCDF. To handle the multiplicity and get a physical meaning for $W(E)$, we concocted an energy scale based on multiples of an energy interval ϵ containing very many levels:

$$(n-1/2)\epsilon < E < (n+1/2)\epsilon$$

Then we can write:

$$W_\epsilon(E) = \sum_{\{E_j\}} \delta_{E, \Sigma E_j} \prod_{j=1}^{N} W_j(E_j)$$

The degeneracies of independent systems do not apply. We have introduced a representation for the Kronecker δ function:

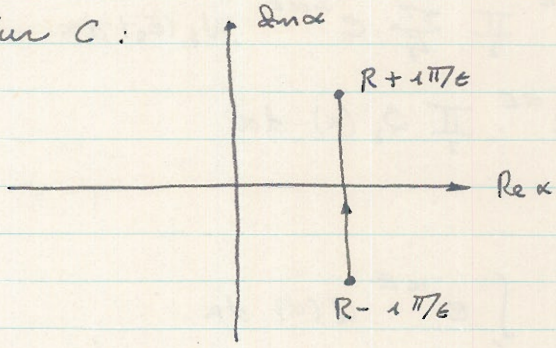
$$W_\epsilon(E) = \frac{\epsilon}{2\pi i} \sum_{\{E_j\}} \int_C e^{\alpha E} \prod_{j=1}^{N} e^{-\alpha E_j} W_j(E_j)\, d\alpha = \frac{\epsilon}{2\pi i} \int_C e^{\alpha E} Q(\alpha)\, d\alpha$$

where: $Q(\alpha) = \prod_{j=1}^{N} Q_j(\alpha)$.

Justification of the interchange of Σ, Π : Take $N=2$:

$$\sum_{E_1, E_2} \prod_{j=1}^{2} F_j(E_j) = \sum_{E_1, E_2} F_1(E_1) F_2(E_2) = \left\{ \sum_{E_1} F_1(E_1) \right\} \left\{ \sum_{E_2} F_2(E_2) \right\}$$

Recall the contour C:



We claim that the integral is not essentially altered by making the limits on $\operatorname{Im}\alpha$ $\pm i\infty$, but if this is so, most of the contribution to the integrand must come from around $\operatorname{Re}\alpha = R$, $\operatorname{Im}\alpha \approx 0$.

We write $W_\epsilon(E)$ as:

$$W_\epsilon(E) = \frac{\epsilon}{2\pi i} \int_C e^{\alpha E + \log Q(\alpha)} d\alpha$$

If we do by the saddle point method, we will find that the greatest contribution comes from around $\mathrm{Re}\,\alpha = R$; $\mathrm{Im}\,\alpha \approx 0$. The saddle point is given by:

$$E = -\frac{d}{d\alpha} \log Q(\alpha)$$

which gives $\alpha = \beta$ for the saddle point. Choose the contour that passes thru the saddle point $(R = \beta)$. Then the only difference between the CM and QM cases is that the limits are different. But when considering the CM case, we found that the great contributions come from around $|\mathrm{Im}\,\alpha| \leq \frac{\beta}{\sqrt{N}}$.

Now in order to replace the terminated contour by the infinite contour we require that:

$$\pi/\epsilon \to \infty \quad \text{if} \quad \pi/\epsilon \gg \frac{\beta}{\sqrt{N}} \quad ; \text{ or, } \quad \epsilon \ll \frac{\pi\sqrt{N}}{\beta}$$

Although we have not really shown that $\beta = \frac{1}{kT}$ we suspect it from our analysis of the perfect gas law. We then use this and claim we can use the infinite contour providing:

$$\epsilon \ll \pi \sqrt{N}\, kT$$

Now it is unthinkable that ϵ would be of the order $\sqrt{N}\, kT$; usually it is of order kT. This still keeps plenty of levels inside ϵ. The errors involved above are exponentially small. Thus the Gaussian correction is very narrow compared with $-\pi/\epsilon$ to π/ϵ and we can make the transition to $\pm\infty$. The result of the integration is, by the same arguments as used in the CM case:

$$\boxed{W_\epsilon(E) = \frac{\epsilon\, e^{\beta E}\, Q(\beta)}{\sqrt{2\pi\, (\log Q(\beta))''}}}$$

The consequences of missing the saddle point and all subsequent corrections are identical with those of the CM treatment.

Proceeding, we recall:

$$P_{E_1}^{(1)} = \frac{W^{(N-1)}(E-E_1)}{W^{(N)}(E)}$$

We now use the fact that missing the saddle point slightly amounts to negligible error, as shown in the CM case. Then:

$$P_{E_1}^{(1)} \doteq \frac{e^{-\beta E_1}}{Q_1(\beta)}$$

This is the same form as obtained in the classical case. We now form the density operator for subsystem 1:

$$\boxed{\rho_1 = \sum_{E_1, k} |E_1 k\rangle P_{E_1}^{(1)} \langle E_1 k|}$$

or, using the asymptotic result:

$$\rho_1 = \sum_{E_1, k} |E_1 k\rangle \frac{e^{-\beta E_1}}{Q_1(\beta)} \langle E_1 k| = \frac{e^{-\beta H}}{Q_1(\beta)} \sum_{E_1 k} |E_1 k\rangle\langle E_1 k|$$

and:

$$\boxed{\rho_1 = \frac{e^{-\beta H}}{Q_1(\beta)}}$$

This is almost identical to the CM case except that now H is an operator. Now recall the definition of the mean value $\overline{\langle A\rangle}$:

$$\overline{\langle A\rangle} = \text{Trace}\, \rho A = \frac{1}{Q(\beta)} \text{Trace}\left\{e^{-\beta H} A\right\} = \frac{1}{Q(\beta)} \text{Trace}\left\{A e^{-\beta H}\right\}$$

Let $A = 1$; and find:

$$\boxed{Q(\beta) = \text{Trace}\, e^{-\beta H}}$$

Choose some set of orthogonal states denoted by $|l\rangle$; then:

$$Q(\beta) = \sum_{l} \langle l | e^{-\beta H} | l \rangle$$

If we had chosen $|l\rangle$ to be the eigenstates of H, then:

$$Q(\beta) = \sum_{E,k} \langle Ek | e^{-\beta H} | Ek \rangle = \sum_{E,k} e^{-\beta E} = \sum_{E} e^{-\beta E}\, W(E)$$

Hence we are back where we started, having now re-obtained the definition of the partition function.

We now digress to consider some of the finer aspects of the Dirac formalism. We take as an example the 1-D QM harmonic oscillator:

$$H = \frac{1}{2m} p^2 + \frac{1}{2} m \omega^2 q^2$$

We know as a result of matrix mechanics and the ladder operator method that we have the following eigenvalue equation and eigenvalue formula:

$$H|n\rangle = E|n\rangle$$

where: $E = \hbar\omega(n + 1/2)$. We can also discuss the matrix elements of the p's and q's using only the algebraic properties of p and q. We now go to a coordinate representation which will give us the ordinary wave mechanics. In this representation we have the operator rules:

$$q|x\rangle = x|x\rangle \quad ; \quad p|x\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}|x\rangle$$

where the states $|x\rangle$ form a complete orthonormal set:

$$\langle x | x' \rangle = \delta(x - x')$$

We use δ function normalization because x, x' are continuous.

The δ function is necessary because the product of the two finite quantities:

$$\int g^*(x')\,\langle x'|\,dx' \cdot \int f(x)\,|x\rangle\,dx$$

cannot be finite itself unless $\langle x|x'\rangle = \delta(x-x')$. We now transform the harmonic oscillator problem to this new representation:

$$\int \langle x|\,\tfrac{1}{2m}\,(p^2+m^2\omega^2 q^2)\,|x'\rangle\langle x'|n\rangle\,dx' = E\,\langle x|n\rangle$$

This is really the Schroedinger equation. In order to see this, consider:

$$\int \langle x|\,\tfrac{1}{2m}\left(-\hbar^2\frac{\partial^2}{\partial x'^2}+m^2\omega^2 x'^2\right)|x'\rangle\langle x'|n\rangle\,dx' = E\,\langle x|n\rangle$$

$$\int \tfrac{1}{2m}\left(-\hbar^2\frac{\partial^2}{\partial x'^2}+m^2\omega^2 x'^2\right)\delta(x-x')\,\langle x'|n\rangle\,dx' = E\,\langle x|n\rangle$$

$$\int \delta(x-x')\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_i^2}+\frac{m\omega^2}{2}x'^2\right)\langle x'|n\rangle\,dx' = E\,\langle x|n\rangle$$

or:

$$\left\{\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{m\omega^2}{2}x^2\right\}\langle x|n\rangle = E\,\langle x|n\rangle$$

We see that this is the Schroedinger equation and the representation functions $\langle x|n\rangle$ are wave functions $\psi_n(x)$. We have then unified matrix mechanics and wave mechanics. The quantity that plays the role of the wave function is the transformation function $\langle x|n\rangle$ which is the Hilbert space analog of direction cosines.

We now return to a discussion of partition functions and the density operator. Recall:

$$Q(\beta) = \text{Trace } e^{-\beta H} = \sum_l \langle l|\,e^{-\beta H}\,|l\rangle$$

There is nothing arbitrary about the trace as might first be thought because it is invariant under a unitary transformation and is the same in all representations.

Form:

$$\text{Trace } e^{-\beta H} = \sum_{l\,\xi'\xi} \langle l|\xi'\rangle\langle\xi'|e^{-\beta H}|\xi\rangle\langle\xi|l\rangle$$

Now, since we have seen that the $\langle\xi|l\rangle$ are wave functions, they must have completeness relations:

$$\sum_l \langle\xi|l\rangle\langle l|\xi'\rangle = \langle\xi|\xi'\rangle = \delta_{\xi\xi'}$$

Then:

$$Q(\beta) = \sum_{\xi\xi'} \delta_{\xi\xi'}\langle\xi'|e^{-\beta H}|\xi\rangle = \sum_{\xi}\langle\xi|e^{-\beta H}|\xi\rangle$$

which shows that $Q(\beta) = \text{Trace } e^{-\beta H}$ is independent of the representation. Let us look at it in the coordinate representation, that is, let $\xi \to x$:

$$Q(\beta) = \sum_l \iint dx\,dx'\,\langle l|x'\rangle\langle x'|e^{-\beta H}|x\rangle\langle x|l\rangle$$

$$= \sum_l \iint dx\,dx'\,\psi_l^*(x')\langle x'|e^{-\beta H}|x\rangle\psi_l(x)$$

Suppose we are considering some general potential $V(x)$, then:

$$Q(\beta) = \sum_l \iint dx\,dx'\,\psi_l^*(x')\,e^{-\beta\left\{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right\}}\delta(x-x')\,\psi_l(x)$$

$$= \sum_l \int dx\,\psi_l^*(x)\,e^{-\beta\left\{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right\}}\psi_l(x)$$

where in the last step we have changed $x' \to x$. We see the invariance again. It, of course, will prove most convenient to take the trace in the eigenstates of the Hamiltonian if these are known and available.

LECTURE 19: 12-5-61

Recall we found for the density operator:

$$\boxed{\rho = \frac{e^{-\beta H}}{Q(\beta)} \quad ; \quad \text{Trace}\, \rho = 1}$$

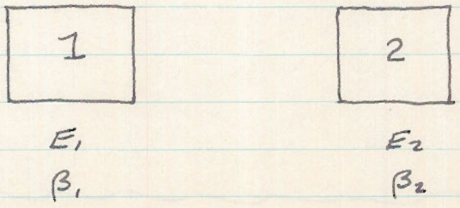
When deriving the above expression, we used the device of dividing up the system into intervals of energy ϵ, each interval containing many levels. Now, this means that, in terms of the assumptions upon which the derivation was based, we cannot treat subsystems of arbitrarily small size, because these subsystems must have many levels within an interval ϵ. For many actual systems, the energy interval ϵ can be about .01 ev, so generally ϵ must be greater than this but less than $\sqrt{3N}\, kT$. Hence, in terms of the derivation, we can only say that ρ applies to a many molecule subsystem. Somewhat of a dilemma exists because we would like to extend the use of ρ to smaller subsystems. We see that ρ no longer contains ϵ explicitly, so we assume that we can still apply ρ to smaller subsystems, even single molecules (paramagnetic susceptibility), although the derivation does not allow this.

Once we have determined ρ, the remainder of SM becomes mostly a matter of integration or summation, so that an understanding of what ρ is comprises the essential problem of SM.

We now go to seriously consider what β is.

Relation Between Statistical Mechanics and Thermodynamics:

Consider two systems, initially independent:

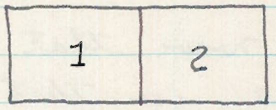


We put each system into the microcanonical distribution, meaning we can specify the energy exactly. β_1 and β_2 can be defined by computing χ_1 and χ_2 for each system and using the definition of β:

$$\{\log \chi_1(\beta_1)\}' = -E_1 \quad \text{defines } \beta_1$$

$$\{\log \chi_2(\beta_2)\}' = -E_2 \quad \text{defines } \beta_2$$

We now put the two systems together and let them interact slightly:



We can no longer tell how the energy is divided, but the total must be the same. Now let the total system come to equilibrium. We have said that the interaction is weak, so we can write the total partition function can be written as a product:

$$\chi = \chi_1 \chi_2$$

Now, when the total compound system reaches equilibrium, we have a new β defined by:

$$\{\log \chi(\beta)\}' = -(E_1 + E_2)$$

or:

$$\{\log \chi_1(\beta)\, \chi_2(\beta)\}' = -(E_1 + E_2)$$

or:

$$\{\log \chi_1(\beta)\}' + \{\log \chi_2(\beta)\}' = \{\log \chi_1(\beta_1)\}' + \{\log \chi_2(\beta_2)\}'$$

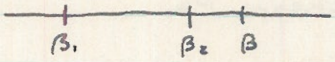
or:

$$\{\log \chi_1(\beta)\}' - \{\log \chi_1(\beta_1)\}' = -\left[\{\log \chi_2(\beta)\}' - \{\log \chi_2(\beta_2)\}'\right]$$

If we take $\beta_1 \geq \beta_2$, and we know that $(\log X)'$ is a monotonically increasing function of β, we can see that the following inequality must be satisfied:

$$\beta_2 \leq \beta \leq \beta_1$$

To see this, suppose β lies outside the interval:



We would have + on the LHS of the above equation in the X's and − on the RHS, and hence a contradiction. For $\beta < \beta_1, \beta_2$, the signs would reverse, again giving a contradiction.

Then suffice it to say now that β has one of the properties of temperature in that two systems at different temperatures initially come to equilibrium at some temperature inbetween when placed in contact. We need not use the microcanonical distribution to do this example, but it could also have been demonstrated using the canonical ensemble where the mean energy is the specified quantity.

We shall now attempt to derive thermodynamics from the principles of Statistical Mechanics. To do this, we must now think of systems capable of undergoing changes in time, that is, the Hamiltonian contains some parameter depending explicitly on time:

$$H = H\left(q_1 \cdots p_F, a(t)\right)$$

Recall that all systems we have talked about were implicitly assumed to be contained in a box as far as the positional coordinates are concerned, that is, it could be that the volume would be a time-dependent parameter. Other examples would be a magnetic field, the force constant of a harmonic oscillator, etc. These parameters introduce macroscopic changes in time, and the statistical statements we make are no longer independent of time.

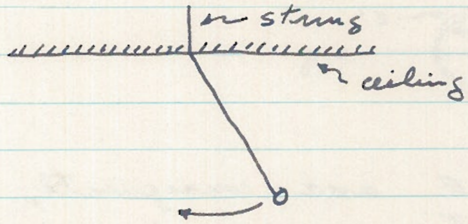
It makes a great deal of difference the way we change these parameters, the detailed way in which they are changed greatly influencing the information we have. This point is neglected in most texts. We will find that the only meaningful mechanical statements that can be made are for infinitely slowly changing systems.

A similar statement underlies the basic principles of thermodynamics, that of reversibility and irreversibility, and we are only able to obtain meaning for reversibility, which says that systems always remain in equilibrium during a change. Such a system is called a quasi-static or slowly changing system. The fact that it makes a difference the way systems are changed provides a large element of parallelinity between thermodynamics and statistical mechanics.

We now discuss an historical example. In the old quantum theory, there existed the Bohr-Sommerfeld quantitization rule:

$$\oint p\,dq = nh$$

Einstein asked what would happen when some of the parameters in this equation were changed with time. He gave as an example a pendulum:

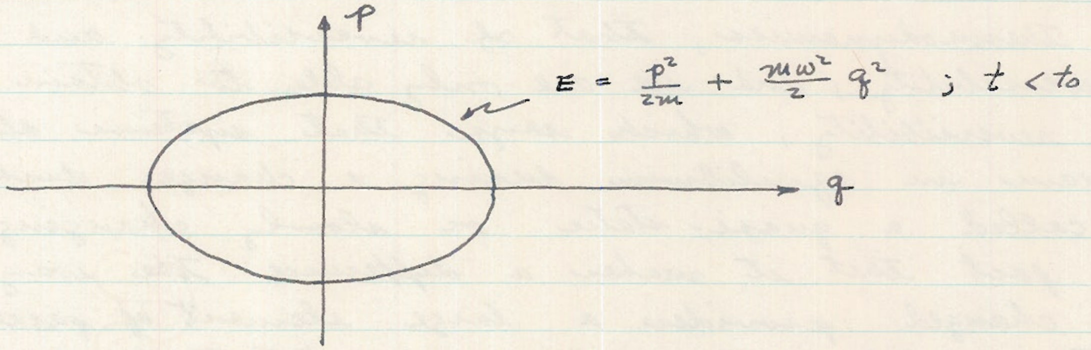


The question is, what happens when we draw the string up into the ceiling while the bob is swinging? What happens to the quantitization? That is, we know $E = n\hbar\omega$ and $\omega^2 = g/l$, but l and hence ω is now changing with time. This was the first attempt to discuss what happens when parameters were changed in the old quantum theory. This problem was first answered by Ehrenfest.

At this time, we will consider what happens when the frequency of a classical harmonic oscillator is changed with time and return to the Q.M. case later.

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 \quad ; \quad \omega = \omega(t)$$

The orbit in phase space for fixed ω for all times, say, $t < t_0$, is:

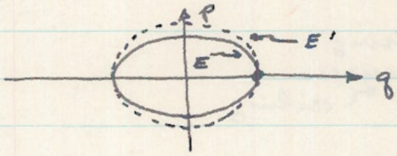


Suppose at $t = t_0$, we instantaneously change $\omega \rightarrow \omega'$. There is not an instantaneous change in p and q, but more of a relaxation effect with a continuous change of canonical variables. What is the final energy? Right after $\omega \rightarrow \omega'$ at $t = t_0$, it is:

$$E' = \frac{p^2(t_0)}{2m} + \frac{m\omega'^2}{2} q^2(t_0) = E + \frac{m}{2}(\omega'^2 - \omega^2) q^2(t_0)$$

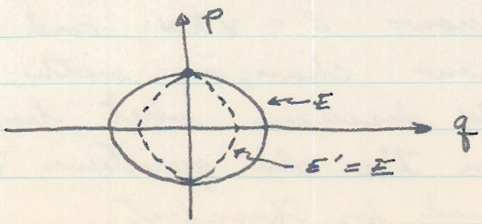
What is the subsequent motion for $t > t_0$? $E' = \frac{p^2}{2m} + \frac{m\omega'^2}{2} q^2$

Suppose at t_0, the oscillator was in the position q_{max}:



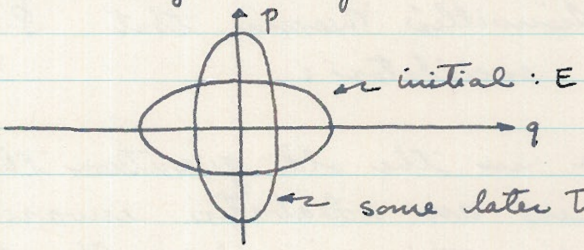
Since, if $\omega' > \omega$, $E' > E$ and consequently the new orbit must cut the momentum axis at a higher point in order to maintain the constant energy E':

For $q(t_0)$ initially 0, we have $E' = E$, but $E' = \frac{p^2}{2m} + \frac{m\omega'^2}{2} q^2$ and $\omega' > \omega$ so the new orbit must cut the q axis at a smaller value:



If we consider a microcanonical ensemble of oscillators with distributed initial conditions over the elliptical orbit, a sudden change in frequency will cause the ellipse to change shape and oscillate around its initial shape. The intermediate states are <u>not</u> ellipses and this provides an example of where a sudden change in a parameter leads to a density function ρ which changes with time.

Now consider the other extreme where the change in ω is made very slowly:



Now, to a good approximation, we the particle completes one orbit, it is back very nearly at its starting point. Hence, at some later time, long compared with the period of the orbit, we have the particle going around another figure but which is still an ellipse. If we consider an entire ensemble, all points are going around the same ellipse at any one time. This is because in one orbit there is only an infinitesimal change in the ellipse. After some time, we have a new ellipse and hence a new energy E'. We do not know E', but we have a way of calculating it since we know that the distribution is still microcanonical. Recall that the Liouville Theorem allows the Hamiltonian to be dependent on time and which says that the volume of phase space enclosed by the system is preserved under a dynamical transformation. Now in our example of slowly changing ω, we have the additional feature of knowing that the shapes of the initial and final orbits are both ellipses. Recall the area of an ellipse:

$$A = \pi \, p_{max} \, q_{max}$$

Then: $A = \pi \sqrt{2mE}\,\frac{1}{\omega}\sqrt{\frac{2E}{m}} = \frac{2\pi E}{\omega}$; $A' = \frac{2\pi E'}{\omega'}$

Now, by Liouville's Theorem, $A = A'$, and we have:

$$\frac{E'}{\omega'} = \frac{E}{\omega}$$

This is an example of an adiabatic Invariant. More generally, we have for periodic systems:

$$\text{Area} = \iint dp\,dq = \oint p\,dq$$

Hence, we see by Liouville's Theorem that $\oint p\,dq$ is invariant. For an oscillator:

$\oint p\,dq \approx \frac{E}{\omega} \sim nh$ in the old quantum theory. Hence we see that $\oint p\,dq$ is an adiabatic invariant under slow changes in ω. This is how Ehrenfest answered Einstein, that is, the string is pulled in at an adiabatic rate so that the Bohr-Sommerfeld quantitization rule is left invariant. Prior to 1925, it was thought that only quantities left under adiabatic changes could be quantized.

What we wanted to illustrate with the above example of the Harmonic Oscillator was that changing the parameters slowly does not disturb the equilibrium while ρ is changing. For ρ initially a MCDF, the final ρ will also be a MCDF over the new ellipse.

The above arguments also extend to multiply-periodic systems. The parameter changes must always be slow compared to the largest period of the system. However, most systems are not periodic and the question arises as to how should we choose the time scale by which to judge adiabatic changes. We cannot tell in general for CM systems, however, some progress can be made with QM systems. Prior to 1925, there were no general rules by which a general mechanical system could be subjected to quantitization. The discovery of the Schroedinger equation removed all these problems. Hence we should expect that time scales can be found for the general QM system by which we can judge adiabatic invariance.

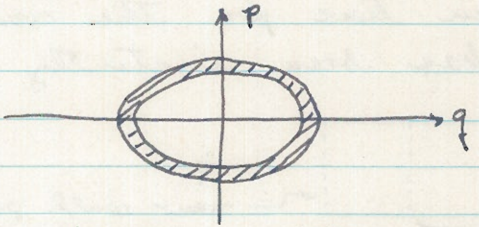
LECTURE 20: 12-7-61

Recapitulation: Sudden changes of a parameter causes the MCDF to change into a function of time. For the harmonic oscillator, the ellipse changes shape with time periodically, and its shape is not generally elliptic while changing.

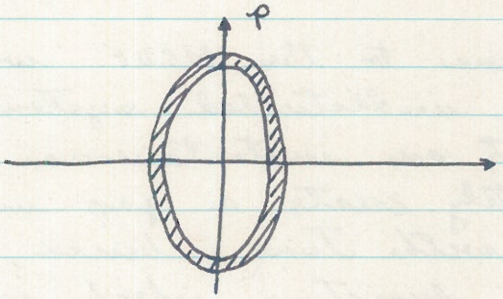
For slow changes, the MCDF remains a MCDF and the initial ellipse of the harmonic oscillator deforms continuously into other ellipses because the orbit in phase space closes almost on itself and the period is well defined.

In order to gain more insight and to actually show that the MCDF is preserved, let us consider some more examples.

Consider first the Harmonic Oscillator again, but this time with some other distribution than MCDF.



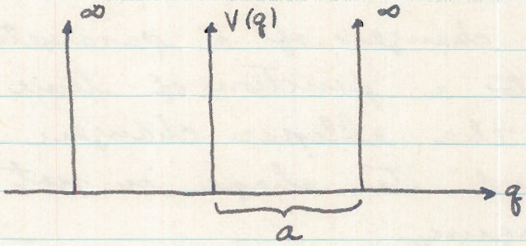
Initial State



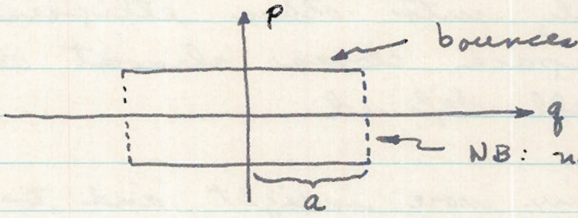
State at later time after the oscillator has undergone adiabatic changes in its parameters

Now the Liouville Theorem states that the annular area or the density in both the initial and later ellipses is the same. When one shrinks the area of the annular ellipse, one gets the MCDF in the limit, hence the MCDF is preserved under adiabatic changes of the parameters.

Consider now the example of a single particle in a 1-D well:



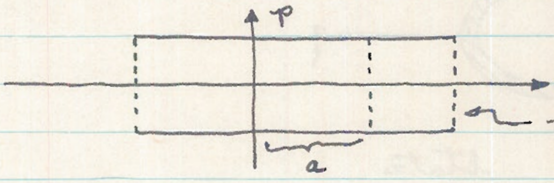
What is the phase space diagram for a definite energy?



bounces off wall, $p \to -p$ instantly.

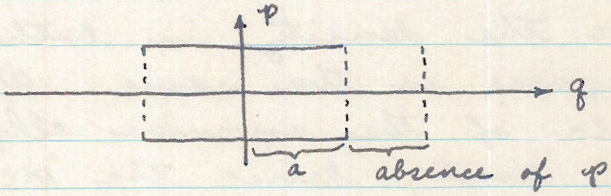
NB: no time spent in this state.

Suppose we change the position of one wall (volume parameter). Consider first the sudden change where the velocity of the wall is much larger than that of the particle. We then have for the new phase space orbit after the wall has been instantly changed:



new wall position

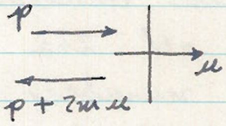
We now ask what happens to the MCDF when the particles are distributed in the undisturbed system as above. It is easy to see that an instantaneous displacement of the wall will immediately create a gap in the distribution which will circulate with time, hence, we have then obtained a distribution function which is changing with time.



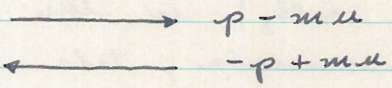
absence of particles

What would happen if we moved the wall in? Take as + velocity, the wall moving outward. Then generally we have:

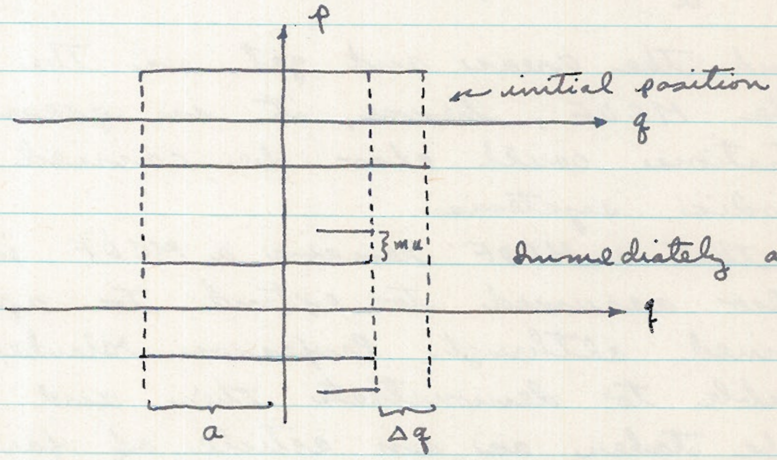
Laboratory System:



Wall Rest System:

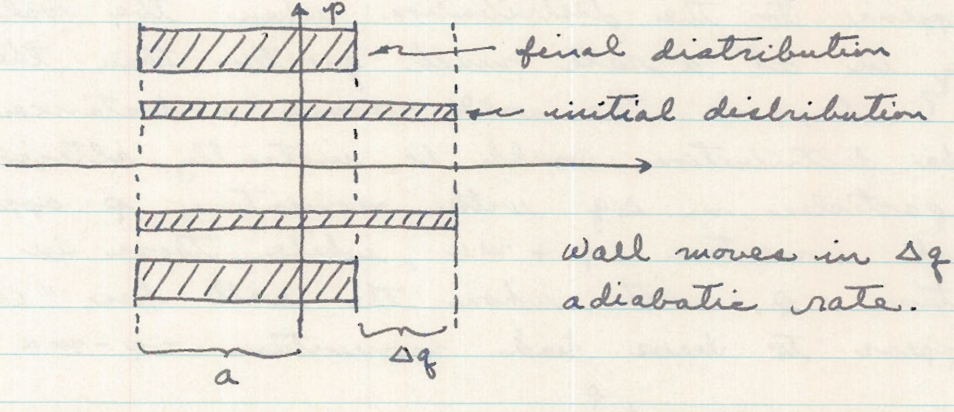


What happens to the distribution when the wall moves suddenly in at a rate much greater than the particle velocity? Suppose the wall moves a distance Δq. Then the distribution would be instantly altered so that those particles in Δq with momentum p appear to have had momentum $p+mu$, while those in Δq with momentum $-p$ with whom the wall has "caught up" now appear to have had momentum $-p-mu$.



We see that we have "split off" the group of particles in Δq and they will now move around their own rectangle independent of the undisturbed particles which now move on the old path but shortened by Δq. Again the MCDF has been made into a time dependent distribution function.

Suppose now we move the wall in slowly compared to the velocity or period of the particle motion and consider that the particle gets an infinitesimal added momentum upon each collision. In an ensemble, we must stipulate that all the particles of the initial ensemble make an equal number of collisions as the wall moves in so the velocity change of all the particles (or each particle, really) is the same and the MCDF remains the MCDF. For example, take some initial distribution and move a wall in on it adiabatically:



Wall moves in Δq at adiabatic rate.

We now shrink the areas and get in the limit for each case the MCDF, hence, it is preserved. The above demonstrations could also be carried out for multiply periodic systems.

The fact that a MCDF remains a MCDF in periodic systems is also assumed to extend to aperiodic systems that are confined although Professor Glauber states that he is not able to demonstrate this and it most likely must be taken as an article of faith. The problem is how does one get an appropriate time scale? One might think of a Poincaré period but one is not usually interested in systems where the time scale that one must vary adiabatically against is the order of gigaparsecs.

Now in QM we get an adiabatic theorem for aperiodic systems. Consider the Hamiltonian $H = H(a)$ where a is a parameter which as yet does not depend upon the time. For example, a could be the frequency in a harmonic oscillator problem.

We then have: $H(a)\,|la\rangle = E_l(a)\,|la\rangle$

Now for fixed a, $|la\rangle$ forms a complete orthonormal set. If we include the time part in a Schroedinger representation, we have:

$$|lat\rangle = e^{-\frac{i}{\hbar}E_l(a)t}\,|la\rangle$$

Now let $a = a(t)$. The Schroedinger equation is:

$$i\hbar\frac{\partial}{\partial t}|t\rangle = H(a(t))\,|t\rangle$$

Now we can expand $|t\rangle$ in a set of stationary states corresponding to the value of a at time t.

$$|t\rangle = \sum_l c_l(t)\,|l, a(t), t\rangle$$

It is important to note that $|l, a(t), t\rangle$ is a stationary state vector (except for $e^{-\frac{i}{\hbar}E(a)t}$). $a(t)$ is taken at time t. The essential time dependence of $|t\rangle$ is contained in the exponential and in $c_l(t)$. Also, we take take $H(a(t))\,|l, a(t), t\rangle = E_l(a(t))\,|l, a(t), t\rangle$. However, we have not specified at which t we are taking $a(t)$ and it could be at each new instant thus allowing us to treat it as a variable of time in the same way $c_l(t)$ is. $H(a(t))\,|l, a(t), t\rangle = E_l(a(t))\,|l, a(t), t\rangle$ apparently seems to require that $a(t)$ vary slowly with time as compared to other quantities like $c_l(t)$ and $e^{-i/\hbar E_l t}$. As a matter of fact, what we assume is that we know the eigenvalues and eigenvectors at each instant of time so we can write a stationary state equation at each instant of time. This is the first step in formulating the adiabatic approximation of time perturbation theory (see Schiff).

Now we could plug $|t\rangle$ back into the Schroedinger equation, but we have no knowledge of the initial conditions so we had better construct the density matrix:

$$\rho = \overline{|t\rangle\langle t|} = \sum_{l,m} |l, a(t), t\rangle\,\overline{c_l(t)\,c_m^*(t)}\,\langle m, a(t), t|$$

We use the Liouville equation:

$$i\hbar \frac{\partial}{\partial t}\rho = -\left[\rho, H(a(t))\right]$$

We simplify the notation by defining:

$$\rho_{\ell m}(t) \equiv \overline{C_\ell(t)\, C_m^*(t)}$$

We will now get a differential equation for $\rho_{\ell m}(t)$, using Liouvilles equation and differentiating everywhere a time dependence appears.

$$i\hbar \sum_{\ell m} \dot a \left\{ \left[\frac{\partial}{\partial a}|\ell, a(t), t\rangle\right] \rho_{\ell m}(t) \langle m, a(t), t| + |\ell, a(t), t\rangle \rho_{\ell m} \frac{\partial}{\partial a}\langle m, a(t), t|\right\}$$

$$+\, i\hbar \sum_{\ell m} |\ell, a(t), t\rangle \frac{\partial \rho_{\ell m}}{\partial t}\langle m, a(t), t| + \sum_{\ell m}(E_\ell - E_m)\,|\ell,a(t),t\rangle\, \rho_{\ell m}\langle m, a(t), t|$$

$$= \sum_{\ell m}(E_\ell - E_m)\,|\ell, a(t), t\rangle\, \rho_{\ell m}\langle m, a(t), t|$$

Hence:

$$\sum_{\ell m}|\ell\rangle \frac{\partial\rho_{\ell m}}{\partial t}\langle m| = -\dot a \sum_{\ell m}\left\{\frac{\partial}{\partial a}|\ell\rangle\, \rho_{\ell m}\langle m| + |\ell\rangle\,\rho_{\ell m}\frac{\partial}{\partial a}\langle m|\right\}$$

Bracket between $\langle \ell_0| \cdots |m_0\rangle$ to form a matrix element. Then we get:

$$\frac{\partial \rho_{\ell m}}{\partial t} = -\dot a \sum_j \left\{\langle \ell|\frac{\partial}{\partial a}|j\rangle\, \rho_{jm} + \rho_{\ell j}\left(\frac{\partial}{\partial a}\langle j|\right)|m\rangle\right\}$$

now define the quantity: $R_{\ell m} = -i\hbar\,\langle \ell, a(t), t|\frac{\partial}{\partial a}|m, a(t), t\rangle$

note that: $\langle \ell|m\rangle = \delta_{\ell m}$, and that:

$$\left(\frac{\partial}{\partial a}\langle \ell|\right)|m\rangle = -\langle \ell|\frac{\partial}{\partial a}|m\rangle, \quad \text{because} \quad \frac{\partial}{\partial a}\delta_{\ell m} = 0.$$

that is: $-i\hbar\left(\frac{\partial}{\partial a}\langle\ell|\right)|m\rangle = i\hbar\,\langle \ell|\frac{\partial}{\partial a}|m\rangle = -R_{\ell m}$

Then:

$$i\hbar\frac{\partial}{\partial t}\rho_{\ell m} = \dot a\sum_j\left\{R_{\ell j}\,\rho_{jm} - \rho_{\ell j}\,R_{jm}\right\}$$

In matrix notation: $i\hbar \frac{\partial}{\partial t} p = \dot{a}(Rp - pR) = \dot{a}[R,p]$
Note that probability is conserved:

$$\text{Trace } \rho = \text{Trace } p \;; \quad i\hbar \frac{\partial}{\partial t} \text{Trace } p = \dot{a} \text{ Trace } [R,p] = 0$$

Hence normalization is preserved.

Now, we cannot solve the above matrix equation in any generality. We can write the differential equation as an integral equation of the Volterra type:

$$p(t) = p(0) + \frac{1}{i\hbar}\int_0^t \dot{a}\,(Rp - pR)\,dt'$$

We will do by successive approximations. The zeroth order is $p(t) = p(0)$. The first approximation is then:

$$p(t) = p(0) + \frac{1}{i\hbar}\int_0^t \dot{a}(t')\left[R(t')\,p(0) - p(0)\,R(t')\right]dt'$$

Before we began the variation of the parameter a, we had equilibrium and hence $p(0)$ is diagonal. Can $p(t)$ have non-diagonal elements? To see, look at the matrix element:

$$p_{\ell m}(t) = p_{\ell m}(0) + \frac{1}{i\hbar}\sum_j \int_0^t \dot{a}(t')\left(R_{\ell j}\,p_{jm} - p_{\ell j}\,R_{jm}\right)dt'$$

Now $p_{\ell m}(0) = p_{\ell\ell}(0)\,\delta_{\ell m}$, so in the first order, we have:

$$p_{\ell m}(t) = p_{\ell\ell}(0)\,\delta_{\ell m} + \frac{1}{i\hbar}\int_0^t \dot{a}(t')\,R_{\ell m}(t')\,dt' \cdot \left(p_{mm}(0) - p_{\ell\ell}(0)\right)$$

We see that for the element $\ell = m$: $p_{\ell\ell}(t) = p_{\ell\ell}(0)$

For the off-diagonal elements, $\ell \neq m$:

$$p_{\ell m}(t) = \frac{1}{i\hbar}\int_0^t \dot{a}(t')\,R_{\ell m}(t')\,dt' \left(p_{mm}(0) - p_{\ell\ell}(0)\right)$$

Now: $R_{\ell m}(t') = \langle \ell, a, t' | \frac{\partial}{\partial a} | m, a, t' \rangle$

We see that this term oscillates with frequency:

$$\omega_{\ell m} = \frac{E_\ell(a) - E_m(a)}{\hbar}$$

now actually, $E_\ell(a)$ and $E_m(a)$ changes because a changes, but we are going to require that the frequency modulation involved is very small compared to $W_{\ell m}$, that is, we say that $W_{\ell m}$ oscillates very fast in regard to the variation in $a(t)$.

Look at the matrix element of $R_{\ell m}$:

$$R_{\ell m} = \langle \ell, a, t | \frac{\partial}{\partial a} | m, a, t \rangle = \langle \ell, a | e^{\frac{i}{\hbar} E_\ell(a) t} \frac{\partial}{\partial a} e^{-\frac{i}{\hbar} E_m(a) t} | m, a \rangle$$

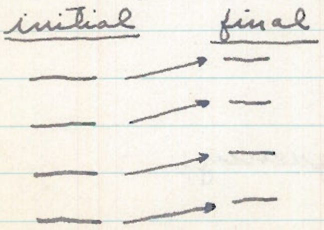
$$= e^{i W_{\ell m} t} \langle \ell a | \frac{\partial}{\partial a} | m a \rangle - \frac{i}{\hbar} \frac{\partial}{\partial a} E_m(a) \, e^{i W_{\ell m} t} \, \delta_{\ell m}$$

But we are interested in the off-diagonal elements, hence, $R_{\ell m} = e^{i W_{\ell m} t} \langle \ell a | \frac{\partial}{\partial a} | m a \rangle$, $\ell \pm m$. Now in the integral $R_{\ell m}$ is highly oscillatory in its time dependence because $e^{i W_{\ell m} t}$ is the controlling factor over a and $\dot{a}$ time dependence. Then we see:

$$\int_0^t \dot{a} \, R_{\ell m} \, dt' \sim \dot{a} \frac{1}{W_{\ell m}}$$

In the limit, we want $\dot{a} \to 0$, $\dot{a}t$ remaining constant, while $t \to \infty$, so that the integral or the off-diagonal elements $\to 0$. We must also stipulate that all derivates of a must be small as well as $\dot{a}$. For example, if we have: $\dot{a} = A(1 + \lambda \sin W_{\ell m} t)$, we see that this is in resonance with $R_{\ell m}$ and we get a finite contribution as $t \to \infty$ even if $A \to 0$ (we get λ). Also we must preclude degeneracy as $W_{\ell m}$ might possibly vanish and a new and different treatment must be used. We can see that the basic unit of the time scale is $\hbar/\Delta E$ and that $\dot{a}$ must be very small or slowly varying with respect to it. Is the first approximation enough? We need to show that second and higher orders go as $\dot{a}^2 t$, at least, and that $\dot{a}^2 t \to 0$ as $\dot{a} \to 0$, $\dot{a}t$ finite, $t \to \infty$.

We see that $\rho_{\ell\ell}(t) = \rho_{\ell\ell}(0)$ or there is no change in the occupancy of the state ℓ during the adiabatic change of a.



A particle remains in the same quantum state with respect to the new parameter at each moment of time. This is a very important result and actually provides the basis for the very existence of equilibrium in the world.

LECTURE 21 : 12-12-61

Recapitulation: We have tried to show that we can talk about stationary states defined at every instant of time while a parameter is undergoing an adiabatic change in time. We have indicated the existence of an adiabatic theorem for QM in that the adiabatic change of a parameter does not change the occupancy or density of a state. (Although it is claimed that no stipulation of periodicity is involved in the QM adiabatic theorem, it seems as though periodicity is necessary for the existence of discreet states, for example, the hydrogenic atom, harmonic oscillator, etc, are periodic systems and provide discreet spectra and hence a time scale $\hbar/\Delta E$ with which one can judge adiabatic cases. It seems as if for a truly aperiodic QM system, for example, the completely free electron, one is just as bad off in the search for a time scale here as one was in the CM aperiodic case because of the continuous nature of the energy spectrum.)

We will not go into any deep questions like how do the degeneracies split when a parameter is varied hence causing a change from the MCDF to something else. From now on, we will always take density functions and operators as remaining stationary under adiabatic changes of the parameters.

We now proceed to consider more fully the relationship between SM and thermodynamics. Consider first a classical system in which energy is not necessarily conserved. The eventual point of the following procedure is to obtain an expression for the change in energy of the system after an adiabatic operation.

$$H = H(q_1 \cdots p_F, a(t))$$

$$\frac{dH}{dt} = \underbrace{\sum_i \left\{ \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i \right\}}_{0,\ \text{because}\ \frac{\partial H}{\partial q_i} = -\dot{p}_i\ ;\ \frac{\partial H}{\partial p_i} = \dot{q}_i} + \frac{\partial H}{\partial a}\frac{da}{dt}$$

Therefore: $\boxed{\frac{dH}{dt} = \frac{\partial H}{\partial a}\frac{da}{dt}}$

For the moment, we will not restrict $\dot{a}$ to be adiabatic so that ρ does not have to be stationary. We have for the mean of H:

$$\overline{H} = \int H(q_1 \cdots p_F, a(t))\, \rho(q_1 \cdots p_F, t)\, \Pi\, dq\, dp$$

We want to compute: $\overline{\frac{dH}{dt}} = \overline{\frac{\partial H}{\partial a}}\frac{da}{dt}$

knowing that $\overline{\frac{dH}{dt}} = \frac{d}{dt}\overline{H}$. Thus:

$$\frac{d}{dt}\overline{H} = \overline{\frac{\partial H}{\partial a}}\frac{da}{dt}$$

Integrating, we have:

$$\boxed{\overline{H}\Big|_t - \overline{H}\Big|_0 = \int_0^t \overline{\frac{\partial H}{\partial a}}\frac{da}{dt'}\, dt'}$$

This is a perfectly general expression, but $\overline{\frac{\partial H}{\partial a}}$ will be hard to compute unless ρ is simple. ρ can be made simple by performing the changes adiabatically, that is, we make $\dot{a}$ slow enough such that:

$$\overline{\frac{\partial H}{\partial a}}\Big|_{t'} = \overline{\frac{\partial H}{\partial a}}\Big|_{a=a(t')} + \mathcal{O}(\dot{a})$$

If a changes slowly, ρ is always static at an instantaneous value of a. If a change of a preserves equilibrium, we need only evaluate things at the instantaneous values of a. Then:

$$\overline{H}\Big|_t - \overline{H}\Big|_0 = \int_0^t \left(\overline{\frac{\partial H}{\partial a}}\Big|_{a=a(t')} + \mathcal{O}(\dot{a}) \right) \dot{a}\, dt'$$

now, in the limit $\dot{a} \to 0$ as $t \to \infty$, $\dot{a}t$ remaining finite, that is, $\int_0^t \dot{a}\, dt = a_1 - a_0$ and is fixed, and the terms in $\dot{a}^2$, $\dot{a}\, \mathcal{O}(\dot{a})$, vanish. We now change variables since the only time dependence left in $\frac{\partial H}{\partial a}$ is in $a(t)$, hence we can change from the variable t to a.

Let: $a = a_0 + \int_0^t \dot{a}\, dt$; $da = \dot{a}\, dt$

Therefore: $\bar{H}\big|_t - \bar{H}\big|_0 = \int_{a_0}^{a_1} \frac{\partial H}{\partial a}\, da$

Note well that the above result is not trivial as it depends on the validity of the adiabatic principle. The time can be eliminated because ρ is stationary.

We now give names to the quantities involved above. If $\bar{H}|_t > \bar{H}|_0$, then the change in energy of the system is positive or there has been negative work done by the system, that is:

$$\bar{H}\big|_t - \bar{H}\big|_0 = -W$$

At this point, we should note that we have restricted all energy changes to changes in the parameter a. We have not considered heat transferral, conduction, or radiation.

Consider for a moment a typical thermodynamic system, that of a fluid in a container. When the container undergoes a change in volume, we know from elementary thermodynamics that the work done is $\delta W = p\, dV$ if the change is done quasistatically. If the piston is pulled out faster than the molecular velocity, no work at all is done, hence we see that $\delta W = p\, dV$ is the maximum possible work done. The total work done is:

$$W = \int_{V_0}^{V_1} p\, dV$$

Note the similarity between this and : $W = -\int_{a_0}^{a_1} \frac{\partial H}{\partial a}\, da$
Also both hold only for quasistatic changes.

motivated by this similarity, we will define a generalized force (not necessarily pressure) associated with a (not necessarily volume):

$$\mathcal{F}_a \equiv -\frac{\partial H}{\partial a}$$

Then, replacing t by a_1 and 0 by a_0:

$$\overline{H}\big|_{a_1} - \overline{H}\big|_{a_0} = -\int_{a_0}^{a_1} \overline{\mathcal{F}_a}\, da = -W \quad \text{, or:}$$

$$\boxed{W = \int_{a_0}^{a_1} \overline{\mathcal{F}_a}\, da}$$

$\mathcal{F}_a$ can be any one of the intensive thermodynamic quantities.

We now discuss the SM and thermodynamic meanings of the term "adiabatic."

SM: a slow change of the parameters of a system with respect to some time scale; a quasistatic process.

Thermodynamics: no heat or energy exchange with the outside world.

In finding the form of W above, we have involved both meanings because we used a stationary ρ and assumed the only energy changes were due to changes in the parameters with no exchange with the outside world. Ehrenfest first used the word "adiabatic" in its SM sense in dealing with a phonon gas.

We now ask for the QM analog for the expression for W. We first determine if:

$$\frac{d}{dt}\overline{\langle G \rangle} = \overline{\left\langle \frac{dG}{dt} \right\rangle}$$

$$\frac{d}{dt}\overline{\langle G \rangle} = \frac{d}{dt}\,\text{Trace}\,\rho G = \text{Trace}\left(\frac{\partial \rho}{\partial t} G + \rho \frac{\partial G}{\partial t}\right)$$

$$\frac{d}{dt}\langle \mathcal{G}\rangle = \text{Trace}\left\{-\frac{1}{i\hbar}[\rho, H]\mathcal{G} + \rho\frac{\partial \mathcal{G}}{\partial t}\right\}$$

now: $\text{Trace}\,[\rho, H]\mathcal{G} = \text{Trace}\left\{\rho H\mathcal{G} - H\rho\mathcal{G}\right\}$

$$= -\text{Trace}\,\rho\,[G, H]$$

because: $\underset{=0}{\text{Trace}\,[\mathcal{G}\rho, H]} = \text{Trace}\,\mathcal{G}[\rho, H] + \text{Trace}\,[G, H]\rho$

Then: $\text{Trace}\,[\rho, H]\mathcal{G} = -\text{Trace}\,\rho\,[G, H]$, and:

$$\frac{d}{dt}\overline{\langle \mathcal{G}\rangle} = \text{Trace}\left\{\rho\left(\frac{1}{i\hbar}[\mathcal{G}, H] + \frac{\partial \mathcal{G}}{\partial t}\right)\right\} = \text{Trace}\left(\rho\frac{d\mathcal{G}}{dt}\right)$$

or:

$$\boxed{\frac{d}{dt}\overline{\langle \mathcal{G}\rangle} = \overline{\left\langle \frac{d\mathcal{G}}{dt}\right\rangle}}$$

and for $\mathcal{G} = H$: $\frac{d}{dt}\overline{\langle H\rangle} = \overline{\left\langle\frac{\partial H}{\partial a}\right\rangle}\frac{da}{dt}$, as before. The argument proceeds the same as for the CM case and we obtain:

$$\mathcal{F}_a = -\frac{\partial H}{\partial a} \quad (\text{now an operator!})$$

$$W = \int_{a_0}^{a_1}\overline{\langle \mathcal{F}_a\rangle}\,da$$

What we will do now is prove or demonstrate a whole spectrum of correspondences between SM and Thermodynamic quantities that hold for systems of a large number of particles.

Consider first the classical system described by the MCDF:

$$\rho = \frac{\delta\left(E - H(q_1\cdots p_F, a)\right)}{\Omega(E, a)}$$

$$\overline{\frac{\partial H}{\partial a}} = \frac{1}{\Omega(E, a)}\int \delta\left(E - H(q_1\cdots p_F, a)\right)\frac{\partial H}{\partial a}\,dV$$

Recall the characteristic function which describes the interior of an energy surface:

$$\mathcal{V}_E(\mathcal{H}) = \int_{\mathcal{H}}^{\infty} \delta(E-E')\,dE' = \begin{cases} 1, & \mathcal{H}<E \\ 0, & \mathcal{H}>E \end{cases}$$

Also:

$$\frac{d}{d\mathcal{H}} \mathcal{V}_E(\mathcal{H}) = -\delta(E-\mathcal{H})$$

Then what is $\frac{\partial}{\partial a} \mathcal{V}_E(H(a))$?

$$\frac{\partial}{\partial a} \mathcal{V}_E(H(a)) = -\delta(E-H)\frac{\partial H}{\partial a}$$

Then:

$$\overline{\frac{\partial H}{\partial a}} = -\frac{1}{\Omega(E,a)} \frac{\partial}{\partial a} \int \mathcal{V}_E(H(a))\,dV = -\frac{1}{\Omega(E,a)} \frac{\partial}{\partial a} \int_{H<E} dV$$

Recall: $V(E,a) = \int_{H<E} dV$; Then:

$$\overline{\frac{\partial H}{\partial a}} = -\frac{1}{\Omega(E,a)} \frac{\partial}{\partial a} V(E,a)$$

or:

$$\boxed{\overline{\mathcal{F}}_a = \frac{1}{\Omega(E,a)} \frac{\partial}{\partial a} V(E,a)}$$

Now consider a change in the phase space volume of the system due to changes in both E and a. This describes a more general mechanical system than we have talked about before as now the change in energy is not only due to the change in a, but something else besides.

$$d\,V(E,a) = \left.\frac{\partial V}{\partial E}\right|_a dE + \left.\frac{\partial V}{\partial a}\right|_E da$$

Recall: $\left.\frac{\partial V}{\partial E}\right|_a = \Omega(E,a)$, then, using $\mathcal{F}_a$ above:

$$d\,V(E,a) = \Omega(E,a)\,dE + \Omega(E,a)\,\overline{\mathcal{F}}_a\,da$$

Then:

$$\boxed{\frac{d\,V(E,a)}{\Omega(E,a)} = dE + \mathcal{F}_a\,da = dE + \delta W}$$

This is entirely analogous to, and in fact equal to, the <u>First Law of Thermodynamics</u>. It is important to recall now that δ means small quantity, not a differential. The usual form of the first law is:

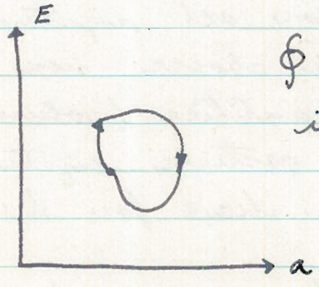
$$\delta Q = dU + \delta W$$

dU is the internal energy and is uniquely defined. On the other hand, δQ, δW may not vanish when taken around a closed loop in its state space. We now can draw the following conclusions from the two forms of the first law:

$$E = U \quad ; \quad \delta Q = \frac{d\,V(E,a)}{\Omega(E,a)}$$

The existence of δQ means the presence of non-adiabatic processes in the thermodynamic sense. We know from the Liouville Theorem that changing a cannot change $V(E,a)$. However, $V(E,a)$ can change by virtue of interaction with the outside world, and so there is a direct correlation between δQ and a change in $V(E,a)$.

$\delta Q = \dfrac{d\,V(E,a)}{\Omega(E,a)}$ finally gives a mechanical interpretation of heat. If it were not for the $\Omega(E,a)$ in the denominator, δQ would be exact. If we integrate δQ around a closed loop in the E-a plane:



$$\oint \frac{1}{\Omega(E,a)}\,d\,V(E,a) \neq 0$$

in general.

We now consider the Second Law of Thermodynamics which states that although δQ is not exact, it has an integrating factor that makes it exact and uniquely defined; and that this integrating factor is $\frac{1}{T}$, such that

$dS = \frac{\delta Q}{T}$, where S is the entropy. The second law states that the integrating factor is the same for all systems in equilibrium with each other. Now, integrating factors are in themselves not unique and if we know one we know the others. That is, $\frac{F'(S)}{T}$ is an integrating factor also because $\frac{F'(S)}{T}\,\delta Q = F'(S)\,dS = dF(S)$. The condition that fixes the integrating factor and makes it unique is that it must be the same for all systems in equilibrium. Because entropies are additive, this forces the choice $F'(S) =$ constant and hence $1/T$ satisfies the requirements.

What is the integrating factor, denoted by λ, for: $\frac{dV}{\Omega} = dE + \delta W$?

$\lambda = \Omega$ is one obvious possibility. However, $\Omega(E)$ is not the same for each system of a number of systems in equilibrium with each other. The Ω's can differ vastly, just from the fact that the volumes of the different systems are different. Professor Glauber indicates that some trouble lies in the too accurate specification the MCDF gives the energy. However, since we now know one λ, we should be able now to find the one we want. Specify in general terms by:

$$\lambda \frac{dV}{\Omega} = d\mathcal{S} \quad ; \quad \lambda = \Omega \frac{d\mathcal{S}(V)}{dV}$$

where $\mathcal{S}$ can be any function of V. However, we want to choose $\mathcal{S}$ such that λ is the same for all systems in equilibrium. Choose $\lambda = k\frac{\Omega}{V}$ as this choice has the property of cancelling all extensive quantities (volumes) hence making λ constant for all free particle systems (k is a proportionality constant). We then have for $\mathcal{S}$:

$$\boxed{\mathcal{S}(V) = k \log V(E,a)}$$

(looks suspiciously like Boltzmann equation)

LECTURE 22 : 12-14-61

We have been discussing the connection between classical SM based on the MCDF and Thermodynamics. We have as the analogy to the First Law of Thermodynamics:

$$\boxed{\frac{dV(E,a)}{\Omega(E,a)} = dE + \overline{F_a}\, da}$$

We make the identification: $\delta Q = \frac{dV}{\Omega}$ and define $\frac{1}{T}$ as the integrating factor of δQ which makes it exact and is the same for all systems in equilibrium. More generally, we may claim that:

$$\lambda = \Omega \frac{d\mathcal{S}(V)}{dV}$$

is an integrating factor that satisfies the requirement that: $\lambda\, \delta Q = d\mathcal{S}(V)$. We want to choose $\mathcal{S}(V)$ so that the integrating factor is very insensitive to the extensive properties of systems in equilibrium, That is, we want to make λ intensive. A way to approach this is to choose:

$$\frac{d\mathcal{S}(V)}{dV} = \frac{\text{constant}}{V} = \frac{k}{V}$$

in an attempt to divide out the extensive quantities. If λ is as insensitive as we hope, then we may write:

$$\lambda = \frac{k\Omega}{V} = \frac{1}{T}$$

We see that λ is independent of N and V (at least for free particles) and depends roughly on the energy per molecule. From:

$$\lambda\, \delta Q = d\mathcal{S}(V) = \lambda \frac{dV}{\Omega} = k \frac{dV}{V}$$, we have:

$$\boxed{\mathcal{S}(V) = k \log V(E,a)}$$

Furthermore: $\frac{1}{T} = k\frac{\Omega}{V} = \frac{\partial}{\partial E} k \log V$

or: $\boxed{\frac{1}{T} = \left(\frac{\partial s}{\partial E}\right)_a}$

We see that $s(V)$ plays the role of the thermodynamic entropy S. In the above relations we really have all we need to establish completely the connection between SM and Thermodynamics, all of the latter following from the above. We can begin to see that SM contains much more than would be suspected from thermodynamics. For example, we will see that SM gives a way of calculating the fluctuations in a Thermodynamic quantity while thermodynamics does not.

Proceeding, we may develop:

$$\overline{F}_{a} = \frac{1}{\Omega}\frac{\partial V(E,a)}{\partial a} = \frac{V}{k\Omega}\frac{k}{V}\frac{\partial V}{\partial a} = \frac{V}{k\Omega}\frac{\partial}{\partial a} k \log V$$

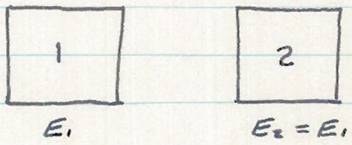
or: $\boxed{\overline{F}_{a} = T\left(\frac{\partial S}{\partial a}\right)_E}$

Note that we have been using E for the internal energy of the system instead of the thermodynamic U. This is because on the MCDF, the internal energy is exactly specified. We shall later modify this.

Elementary texts say that the entropy S is k times the logarithm of the p.s. volume occupied by the probability distribution function. Here we have $V(E,a)$ or the volume inside a constant E surface in p.s., and the volume occupied by the MCDF is zero. However, recall that most of the volume of a hypersphere tends to collect near the surface, hence a slight uncertainty in the energy does not matter much.

All of the above SM relations are of a very general nature in that they make few assumptions about the details of the system. However, when working with the MCDF, the too accurate specification of the energy is found to be not practical.

To examine this matter further, take two systems with identical and precisely specified energies and let them interact slightly:



$E = 2E_1$

Now at some later time the two systems will come together in equilibrium with the result that there will be no way of finding how the energy is divided between 1 and 2. That is, if we insist on taking equilibrium to be when both systems have the same energy, it will mean that $\lambda = \frac{1}{T}$ will not be completely free or independent of the details of the system. That is, initially:

$$\lambda_1 = \frac{1}{T_1} = k \frac{\Omega_1(E_1)}{V_1(E_1)}$$

Finally: $$\lambda = \frac{1}{T} = k \frac{\Omega(E)}{V(E)}$$

Now it will be very nearly true that $\frac{\Omega_1(E_1)}{V_1(E_1)} = \frac{\Omega(2E_1)}{V(2E_1)}$ but not precisely, only asymptotically (to an excellent degree, however) for large N. We would like a definition of equilibrium so that systems do not have to have the same energy to be in equilibrium. We would also like the entropies to be additive: Here we have:

$$S_1 = k \log V_1(E_1) \quad ; \quad S = k \log V(E)$$

However: $\log V(2E_1) \neq 2 \log V_1(E_1)$ because $V(2E_1) \neq V_1^2(E_1)$

The main difficulty is that the MCDF is much too accurate for reality. We begin to depart from this accuracy by approximating $V(E,a)$. Recall the relation:

$$\boxed{V(E,a) = \int_0^E \Omega(E',a)\, dE'}$$

Since most of $V(E,a)$ is near the constant E surface, we need not hesitate to use an approximation that puts E' outside the constant E surface.

Recall the approximation for Ω:

$$\Omega(E') = \frac{e^{\beta E'} \chi(\beta)}{\sqrt{2\pi(\log\chi)''}} \; e^{-\frac{(E'-E)^2}{2(\log\chi)''}}$$

where $E = -\{\log\chi(\beta)\}'$

Remember that we have shown that we can remove the exponential correction if:

$$\left|\frac{E'-E}{E}\right| \leq \mathcal{O}\left(\frac{1}{N^{1/3}}\right)$$ or if E' is near or on the E surface.

However, we retain it for a while. Continuing:

$$V = \frac{e^{\beta E} \chi(\beta)}{\sqrt{2\pi(\log\chi)''}} \int_0^E e^{\beta(E'-E) - \frac{(E'-E)^2}{2(\log\chi)''}} \, dE'$$

Let $x = \frac{E-E'}{E}$, then the integral becomes:

$$\int_0^1 e^{-\beta E x - \frac{E^2}{2(\log\chi)''} x^2} \, dx$$

We are now in a position to examine the orders of magnitude involved. We call on the free gas model as an aid. Recall some figures for this model:

$$\frac{1}{\beta} \sim \frac{E}{N} \; ; \; E = (\log\chi)' \; ; \; (\log\chi)'' \sim \frac{E}{\beta} \sim \frac{E^2}{N}$$

Hence the integral is like:

$$\int_0^1 e^{-\mathcal{O}(N)x - \mathcal{O}(N)x^2} \, dx$$

But: $e^{-\mathcal{O}(N)x} \to 0$ as $x \to 1$ providing $x \gg \frac{1}{N}$

$e^{-\mathcal{O}(N)x^2} \to 0$ as $x \to 1$; $x \gg \frac{1}{\sqrt{N}}$

We can then replace the 1 by ∞ and expand the x^2 term.

We obtain:

$$\int_0^\infty e^{-\Theta(N)x}\left\{1-\Theta(N)x^2\right\}dx$$

$$\to \int_0^\infty e^{-\Theta(N)x}\,dx\left\{1+\Theta\left(\tfrac{1}{N}\right)\right\}$$

so we can drop the Gaussian correction as expected. The approximation to V is then:

$$V(E,a) \doteq \frac{e^{\beta E}}{\beta}\frac{\chi(\beta)}{\sqrt{2\pi(\log\chi)''}} = \frac{\Omega(E,a)}{\beta}$$

Now this corresponds to the statements made about the entropy usually made in the elementary texts. Hence

$$\log V(E,a) = \log\Omega(E,a) - \log\beta + \cdots$$

Now as the system gets large, V increases, Ω increases, but β, being intensive, remains the same. Hence $\log\beta$ can be dropped with respect to $\log\Omega(E\,a)$ because it is so small and we can generally write:

$$\log V(E,a) = \log\Omega(E,a)\left\{1+\Theta\left(\tfrac{1}{N}\right)\right\}$$

Now it is $\log\Omega$ that is found in elementary work and we write to $\Theta\left(\frac{1}{N}\right)$:

$$\frac{1}{T} = \frac{\partial}{\partial E}k\log V = \frac{\partial}{\partial E}k\log\Omega$$

$$\boxed{S = k\log\Omega}$$

Do the thermodynamic relations still hold?

$$\mathcal{F}_a = \frac{1}{\Omega}\frac{\partial V}{\partial a} = \frac{V}{\Omega}\frac{\partial}{\partial a}\log V = T\frac{\partial}{\partial a}k\log V = T\frac{\partial}{\partial a}k\log\Omega$$

$$\therefore\ \mathcal{F}_a = T\frac{\partial S}{\partial a}$$

and the proper relationships are still preserved.

However, we still have no additivity of entropy since we have $\Omega(2E_1) \neq \Omega^2(E_1)$ except logarithmically or asymptotically. We hence try throwing away some more small components of $\Omega(E)$, in particular the normalization. We are beginning to see that SM is much more of an exact science than thermodynamics.

Recall:

$$V(E,a) \approx \frac{\Omega(E,a)}{\beta} = \frac{e^{\beta E}}{\beta} \frac{\chi(\beta)}{\sqrt{2\pi (\log \chi)''}}$$

Then:

$$\log V(E,a) \approx \log \Omega(E,a) \approx \log \left(e^{\beta E} \chi(\beta) \right) \left(1 + O\left(\tfrac{1}{N}\right) \right)$$

neglecting the normalization terms. Hence:

$$\log V(E,a) \approx \beta E + \log \chi(\beta, a)$$

and:

$$\boxed{S = k \left\{ \beta E + \log \chi(\beta, a) \right\}}$$

Form:

$$\frac{1}{T} = k \frac{\partial}{\partial E} \left\{ \beta E + \log \chi(\beta, a) \right\}$$

Now for a fixed, knowing that β is implicitly dependent on E or $\beta = \beta(E,a)$ since $E = -\{\log \chi(\beta, a)\}'$, we have:

$$\frac{1}{T} = k \left(\beta + E \frac{\partial \beta}{\partial E} + (\log \chi)' \frac{\partial \beta}{\partial E} \right) = k \left(\beta + E \frac{\partial \beta}{\partial E} - E \frac{\partial \beta}{\partial E} \right)$$

$$= k\beta$$

At long last we have obtained the formal identification of β:

$$\boxed{\boxed{\beta = \frac{1}{kT}}}$$

All the thermodynamic relations hold as before plus the fact that we now can establish a connection between $\chi(\beta)$ and the thermodynamic quantities.

That is:

$$\overline{\mathcal{F}_a} = \frac{1}{\Omega}\frac{\partial V}{\partial a} = \frac{V}{\Omega}\frac{\partial}{\partial a}\log V = \frac{1}{\beta}\frac{\partial}{\partial a}\log \Omega$$

$$= \frac{1}{\beta}\frac{\partial}{\partial a}\left(\beta E + \log \chi\right)$$

$$= \frac{1}{\beta}\left\{ \frac{\partial \beta}{\partial a}\Big|_E E + \underbrace{\frac{\partial}{\partial \beta}\log \chi}_{-E}\ \frac{\partial \beta}{\partial a}\Big|_E + \frac{\partial}{\partial a}(\log \chi)\Big|_\beta \right\}$$

$$= \frac{1}{\beta}\frac{\partial}{\partial a}(\log \chi)\Big|_\beta$$

or:

$$\boxed{\overline{\mathcal{F}_a} = \frac{1}{\beta}\frac{\partial}{\partial a}\log \chi(\beta,a)\Big|_\beta = kT\frac{\partial}{\partial a}\log \chi(\beta,a)\Big|_T}$$

Note that now we have additivity of entropies:

$$S = S_1 + S_2 = k\left\{\beta \Sigma E_i + \log \chi_1(\beta,a)\chi_2(\beta,a)\right\}$$

$$= k\left\{\beta E_1 + \beta E_2 + \log \chi_1(\beta,a) + \log \chi_2(\beta,a)\right\}$$

We have now arrived at an equivalence with thermodynamics, at least on a mathematical basis, that has no remaining ambiguities.

What about this equivalence in Quantum SM? If we begin with the MCDF, an equivalence cannot be worked out without much difficulty and travail. It is easier to develop from the standpoint of the CDF. However, there are situations where we want to use QSM but the principles of the CDF do not apply. Examples arise in collisions of nuclear systems. Here the MCDF must be used since we specify the energy of the bombarding particle.

What correspondence can we draw for $V(E,a)$, the volume inside a constant energy surface in p.s., seeing that p.s. has no meaning in QM?

The natural analogy is the number of levels below a certain energy E. Thus:

$$V(E,a) \leftrightarrow N(E,a)$$

and: $$dN(E,a) = \frac{\partial N}{\partial E}\,dE + \frac{\partial N}{\partial a}\,da$$

We now recall from previous developments:

$$\frac{\partial N}{\partial E} = \text{level density} \equiv \omega(E,a) = \frac{W_\epsilon(E,a)}{\epsilon}$$

Now, on the MCDF we state:

$$\overline{\langle \mathcal{F}_a \rangle} = \frac{1}{W_\epsilon}\frac{\partial}{\partial a}\int_E^{E+\epsilon} \text{Trace}\ \psi_{E'}\left(\mathcal{H}(a)\right) dE'$$

In our correspondence, we may write:

$$\overline{\langle \mathcal{F}_a \rangle} = \frac{1}{W_\epsilon}\frac{\partial}{\partial a}\,\epsilon N(E,a) = \frac{\epsilon}{W_\epsilon}\frac{\partial}{\partial a}N(E,a) = \frac{1}{\omega}\frac{\partial N}{\partial a}$$

so we have the QSM form of the First Law of Thermodynamics:

$$dN = \omega\, dE + \omega\, \overline{\langle \mathcal{F}_a \rangle}\, da$$

from which we identify as the heat in QSM:

$$\delta Q = \frac{dN(E,a)}{\omega(E,a)}$$

In complete analogy with the classical procedure, we obtain

$$\boxed{S = k \log N(E,a)}$$

$$\boxed{\frac{1}{T} = k\frac{\omega}{N} = k\frac{\partial}{\partial E}\log N}$$

Furthermore:

$$\boxed{\overline{\langle \mathcal{F}_a \rangle} = \frac{1}{\omega}\frac{\partial N}{\partial a} = \frac{N}{k\omega}\frac{\partial}{\partial a} k \log N = kT\frac{\partial}{\partial a}\log N(E,a)}$$

In complete accord with the classical procedure, we have in the next level of approximation:

$$\log N = \log \omega - \log \beta + \cdots$$

from whence:

$$\boxed{\begin{aligned} S &= k \log \omega(E,a) \\ \frac{1}{T} &= k\frac{\partial}{\partial E}\log \omega \\ \overline{\langle \mathcal{F}_a \rangle} &= kT\frac{\partial}{\partial a}\log \omega \end{aligned}}$$

To the third level of approximation:

$$\omega \approx e^{\beta E}\, Q(\beta,a)$$

and:

$$\boxed{\begin{aligned} S &= k\left\{\beta E + \log Q(\beta,a)\right\} \\ \beta &= \frac{1}{kT} \\ \overline{\langle \mathcal{F}_a \rangle} &= kT\frac{\partial}{\partial a}\log Q(\beta,a) \end{aligned}}$$

LECTURE 23: 12-19-61

Recapitulation:

Levels of approximation of SM to thermodynamics:

$$\boxed{\text{First Level:} \quad S = k \log V(E,a)}$$

This is the most general correspondence, regardless of interaction. However, the entropies are not additive here.

$$\boxed{\text{Second Level:} \quad S = k \log \Omega(E,a)}$$

We assumed non-interaction in order to gain an idea of the orders of magnitude involved.

For the third level, from:

$$\Omega = \frac{e^{\beta E} \chi(\beta,a)}{\sqrt{2\pi (\log \chi)''}}$$

we obtain: $S = k \left\{ \underbrace{\beta E}_{O(N)} + \underbrace{\log \chi}_{O(N)} - \underbrace{\tfrac{1}{2} \log 2\pi (\log \chi)''}_{O(\log N)} \right\}$

We can easily drop the 3rd term because $O(N) \sim 10^{23}$ and $O(\log N) \approx 23$. Hence:

$$\boxed{\text{Third Level:} \quad S = k \left\{ \beta E + \log \chi \right\}}$$

At this level, the entropies add and we obtain complete equivalence to thermodynamics, and we identify β with $1/kT$ and as T is the same for all systems in equilibrium, β characterizes equilibrium in the third level of approximation.

One of the difficulties of basing any equivalence on the MCDF is that we have to specify the energy much more accurately than is required by thermodynamics.

A more suitable distribution to use is the CDF as we need not specify E exactly and indeed almost always cannot specify it exactly. Although this approach is as versatile as the MCDF, it is simple, uncomplicated, and directly applicable to QM.

Recall,

$$\overline{\mathcal{F}_a} = -\overline{\frac{\partial H(a)}{\partial a}}$$

Take an ensemble of non-interacting subsystems: $H = \sum H_j$

Then: $$\overline{\mathcal{F}_a} = -\sum_j \overline{\frac{\partial H_j}{\partial a}}$$

We will suppose that we are using the MCDF for the whole system, but the CDF for the subsystem: Then:

$$\overline{\frac{\partial H_j}{\partial a}} = \int \frac{\partial H_j}{\partial a}\rho_j\,dV_j = \frac{1}{\chi_j(\beta,a)}\int \frac{\partial H_j}{\partial a}e^{-\beta H_j}\,dV_j$$

$$= \frac{1}{\chi_j(\beta,a)}\left(-\frac{1}{\beta}\right)\frac{\partial}{\partial a}\chi_j(\beta,a) = -\frac{1}{\beta}\frac{\partial}{\partial a}\log\chi_j(\beta,a)$$

Then, we get a result similar to the MCDF:

$$\boxed{\overline{\mathcal{F}_a} = \sum_j \frac{1}{\beta}\frac{\partial}{\partial a}\log\chi_j(\beta,a) = \frac{1}{\beta}\frac{\partial}{\partial a}\log\chi(\beta,a)}$$

To find the entropy relation, consider the work done by varying all of the parameters of the system, that is:

$$\delta W = \sum_e \overline{\mathcal{F}_{a_e}}\,da_e = \frac{1}{\beta}\sum_e \frac{\partial}{\partial a_e}\log\chi(\beta,a,\cdots)\,da_e$$

Now:

$$\sum_e \frac{\partial}{\partial a_e}\log\chi(\beta,a,\cdots)\,da_e + \frac{\partial}{\partial\beta}\log\chi\,d\beta = d\left\{\log\chi(\beta,a,\cdots)\right\}$$

$$\delta W = \frac{1}{\beta}\left[d\log\chi(\beta,a,\cdots) - \frac{\partial}{\partial\beta}\log\chi\,d\beta\right]$$

Use the first law for the whole system which means we specify energy exactly on the MCDF:

$$\delta Q = dE + \delta W \quad ; \quad E = -\frac{\partial}{\partial\beta}\log\chi$$

Then we have: $\delta Q = dE + \frac{1}{\beta}\left[d\log\chi + E\,d\beta \right]$

$$= \frac{1}{\beta}\left[\beta\,dE + E\,d\beta + d\log\chi \right]$$

$$= \frac{1}{\beta}\left[d(\beta E) + d\log\chi \right]$$

It is clear that we have an integrating factor for δQ, namely β. However, for all systems in equilibrium β is the same. Then:

$$\beta\,\delta Q = \frac{dS}{k} \;;\quad \beta = \frac{1}{kT} \text{ from } k\beta = \frac{1}{T}$$

We then finally have:

$$\boxed{S = k\left[\beta E + \log\chi(\beta, a, \cdots) \right]}$$

so we get the same result whether or not we start from the canonical distribution for the subsystem or from the MCDF and approximate down to the first level.

Problem ⑩:

a) Given N particles of mass m in volume v and the partition function for one particle being:

$$\xi(\beta, v) = \left(\frac{2\pi m}{\beta}\right)^{3/2} v$$

Find the equation of state (pressure).

b) Find expressions for S at our 3 levels of approximation and indicate how greatly they differ.

The above development basing the system on the MCDF and the subsystems on the CDF and showing that identical results are obtained is somewhat hampered by the fact that the subsystems must be interacting. This is particularly bothersome when working with solids where all parts usually interact.

what do we do with systems whose parts interact? We could use $S = k \log V(E,a)$, but in general one cannot find $V(E,a)$.

One thing we can do (and a very realistic thing, too) is to specify the state of the system by its temperature instead of energy as we usually do in the laboratory when dealing with large number of things, in equilibrium with their surroundings. We then talk of the system under study as a member of a supersystem (say, the universe whose energy we assume is precisely determined and is describable by the MCDF) which is non-interacting with other members of the supersystem so that the system under study can be described by the CDF regardless of how large or complicated this system may be.

With this notion in mind, we go about transforming our knowledge of independent subsystems to cover the usual case of a single large system in thermal equilibrium with its surroundings. We then write for the distribution function the CDF for specified temperature but unclear energy:

$$\rho = \frac{e^{-\beta H}}{\chi(\beta)}$$

Under these conditions we must now talk about the internal energy in terms of an average. That is, we talk about U and not E now. We then write for the thermodynamic internal energy:

$$\boxed{U = \bar{E} = \frac{1}{\chi(\beta)} \int H e^{-\beta H} dV = -\frac{\partial}{\partial \beta} \log \chi(\beta, a, \cdots)}$$

The fluctuations that will occur in U determines the validity of the application of thermodynamics. The form of the other relations holds:

$$\overline{F_a} = -\overline{\frac{\partial H}{\partial a}} = \frac{1}{\beta} \frac{\partial}{\partial a} \log \chi$$

$$\delta Q = dU + \delta W = \frac{1}{\beta} d(\beta U + \log \chi) \;; \quad \beta = \frac{1}{kT}$$

all thermodynamic quantities have their SM analogs and all thermodynamic problems can be worked out once one knows the partition function χ.

We now indicate how this comes about by deriving some useful relations between thermodynamic quantities and the partition function.

Helmholtz Free Energy

The Helmholtz Free Energy measures the ability of a system to do work isothermally. It is defined as:

$$\boxed{F = U - TS}$$

To see the physical meaning, compute the work done along an isothermal path:

$$W = \int_1^2 \delta W = \int_1^2 (-dU + TdS)$$

$$= -[U]_1^2 + T[S]_1^2 = (U_1 - TS_1) - (U_2 - TS_2)$$

$$= -(F_2 - F_1)$$

so the physical meaning of the free energy F becomes clear. The isothermal path implies the presence of a heat reservoir which maintains the temperature throughout the change, quite suited to our problem. Now, in a quasi-static isothermal process:

$$\delta W \underset{\substack{\uparrow \\ \text{quasistatic}}}{=} p\,dv \underset{\substack{\uparrow \\ \text{isothermal}}}{=} -dF$$

Then: $\boxed{p = -\left.\frac{\partial F}{\partial v}\right|_T}$

We now use the statistical mechanical relation for the entropy:

$$S = k\{\beta U + \log \chi\}$$

Plugging in F:

$$F = U - TS = U - Tk\left\{\frac{U}{kT} + \log Z\right\}$$

or: $$\boxed{F = -kT \log Z(\beta, a, \cdots) = -\frac{1}{\beta}\log Z(\beta, a, \cdots)}$$

This expression is widely used because of its simple relation to SM and is good for problems where T and V are the independent variables. This may be considered the relation that actually links thermodynamics and SM because so many quantities can be found in terms of F. For example consider:

$$dF = dU - d(TS) = \delta Q - \delta W - d(TS)$$

$$= TdS - p\,dv - d(TS) = -S\,dT - p\,dv$$

where T and v are the independent variables. Hence:

$$\boxed{S = -\frac{\partial F}{\partial T}\Big|_v \quad ; \quad p = -\frac{\partial F}{\partial v}\Big|_T}$$

Of course these same relations can be derived through SM and using the above expression for F, $F = -kT \log Z$:

$$\frac{S}{k} = \beta U + \log Z = -\beta\frac{\partial}{\partial\beta}\log Z + \log Z = -\beta^2\frac{\partial}{\partial\beta}\left(\frac{\log Z}{\beta}\right)$$

or: $$S = k\beta^2\frac{\partial}{\partial\beta}\left\{-kT\log Z\right\} = -\frac{\partial}{\partial T}\left\{-kT\log Z\right\} = -\frac{\partial F}{\partial T}$$

where we use: $\beta = \frac{1}{kT}$; $d\beta = -\frac{1}{kT^2}dT = -\beta^2 d(kT)$

Assuming that we have $F = -kT\log Z$, can we find U?

Thermodynamics: $U = F + TS = F - T\frac{\partial F}{\partial T} = -T^2\frac{\partial}{\partial T}\left(\frac{F}{T}\right)$

SM: $U = -\frac{\partial}{\partial\beta}\log Z = \frac{\partial}{\partial\beta}\left(-\frac{F}{kT}\right) = -T^2\frac{\partial}{\partial T}\left(\frac{F}{T}\right)$

Hence, as long as $F = -kT\log Z$, the correspondence between SM and thermodynamics seems complete and consistent.

Heat Capacity : C

We will talk only about the Heat Capacity at constant intensive quantity a. Its definition is:

$$\boxed{C_a \equiv \left.\frac{\partial U}{\partial T}\right|_a}$$

Now: $C_a = \left.\frac{\partial U}{\partial T}\right|_a = \frac{\partial}{\partial T}\left(F - T\frac{\partial F}{\partial T}\right)$

$$= \frac{\partial F}{\partial T} - \frac{\partial F}{\partial T} - T\frac{\partial^2 F}{\partial T^2}$$

or: $$\boxed{C_a = -T\frac{\partial^2 F}{\partial T^2}}$$

Another approach using the internal energy expression:

$$C_a = \frac{\partial}{\partial T}\left\{-T^2\frac{\partial}{\partial T}\left(\frac{F}{T}\right)\right\} = -k\beta^2\frac{\partial^2}{\partial\beta^2}(\beta F)$$

Now $\beta F = \frac{F}{kT} = -\log \mathcal{Z}$ and then:

$$\boxed{C_a = k\beta^2\frac{\partial^2}{\partial\beta^2}\log \mathcal{Z}}$$

More directly: $U = -\frac{\partial}{\partial\beta}\log \mathcal{Z}$

Then: $C_a = \frac{\partial U}{\partial T} = -\frac{1}{T^2}\frac{\partial}{\partial\left(\frac{1}{T}\right)}U = -k\beta^2\frac{\partial}{\partial\beta}U = k\beta^2\frac{\partial^2}{\partial\beta^2}\log \mathcal{Z}$

Now recall that we have shown $\frac{\partial^2}{\partial\beta^2}\log \mathcal{Z}$ is always positive so we see that the statistical mechanics yields a proof for which Thermodynamics provides none, namely:

The Heat Capacity C_a is always positive.

LECTURE 24 : 1-4-62

We have shown that $C_a \geq 0$, because $\frac{\partial^2}{\partial\beta^2}\log \chi \geq 0$. Another way to show this is to consider the fluctuations in the energy u:

$$\overline{(E-\bar{E})^2} = \overline{(E-u)^2} = \overline{E^2} - \bar{E}^2 = \frac{\partial^2}{\partial\beta^2}\log\chi \geq 0$$

Also:

$$\boxed{\overline{(E-u)^2} = \frac{C_a}{k\beta^2} = kT^2 C_a}$$

This relation could not have been obtained by thermodynamics alone. $\overline{(E-u)^2}$ gives error involved in thermodynamics. How great is a typical fluctuation? Consider the perfect gas whose single particle partition function is:

$$\xi(\beta,v) = v\int_{-\infty}^{\infty} e^{-\beta\frac{p^2}{2m}}\,d\vec{p} = \left(\frac{2\pi m}{\beta}\right)^{3/2} v$$

Here $\chi = \xi^N$ for N particles in the gas.

$$\log\chi = N\left\{\log v + \tfrac{3}{2}\log 2\pi m - \tfrac{3}{2}\log\beta\right\}$$

$$u = -\frac{\partial}{\partial\beta}\log\chi = \frac{3N}{2\beta} = \frac{3}{2}NkT$$

$$C_a = \frac{3}{2}Nk$$

Then: $\overline{(E-u)^2} = kT^2\cdot\frac{3}{2}Nk = \frac{3}{2}N(kT)^2$

now take the ratio of the RMS fluctuation to u:

$$\frac{\sqrt{\overline{(E-u)^2}}}{u} = \sqrt{\frac{2}{3N}}$$

This result is quite typical, that is, absolute fluctuations usually go as $\frac{1}{\sqrt{N}} \sim 10^{-11}$ so that we have so small an error involved in thermodynamics that we can usually forget about it.

We have more generally for the fluctuations:

$$\frac{\sqrt{\overline{(E-U)^2}}}{U} = \frac{\sqrt{kT^2 C_a}}{U}$$

Now in general C_a is an extensive quantity and goes as N as does U, so that the fluctuations do go as $1/\sqrt{N}$ times a constant independent of N.

<u>Postscript to the Equipartition Theorem</u>

Recall the definition of the θ-function:

$$\theta(\mathcal{H}) = \int_{\mathcal{H}}^{\infty} \rho(E)\, dE$$

We then found that:

$$\overline{p_j \frac{\partial H}{\partial p_j}} = \overline{q_k \frac{\partial H}{\partial q_k}} = \int \theta(\mathcal{H})\, dV$$

Let us look at the consequences of this theorem in both the MCDF and the CDF:

1) MCDF: $\rho(\mathcal{H}) = \dfrac{\delta(E-\mathcal{H})}{\Omega(E)}$

Then: $$\theta = \frac{1}{\Omega}\int_{\mathcal{H}}^{\infty} \delta(E-E')\, dE' = \frac{\mathcal{V}_E(\mathcal{H})}{\Omega(E)}$$

and: $$\int \theta(\mathcal{H})\, dV = \frac{1}{\Omega(E)}\int \mathcal{V}_E(\mathcal{H})\, dV = \frac{V(E)}{\Omega(E)} = \frac{1}{\frac{1}{V(E)}\frac{dV(E)}{dE}}$$

$$= \frac{1}{\frac{d}{dE}\log V(E)}$$

We now recall: $\dfrac{1}{kT} = \dfrac{d}{dE}\log V(E) = \dfrac{\Omega}{V}$

and hence: $$\boxed{\int \theta(\mathcal{H})\, dV = kT\,; \quad \overline{p_j \frac{\partial H}{\partial p_j}} = \overline{q_k \frac{\partial H}{\partial q_k}} = kT}$$

Thus, in the MCDF, we have the exact well-known relation of the Equipartition Theorem.

2) CDF : $\rho(\mathcal{H}) = \dfrac{e^{-\beta \mathcal{H}}}{\chi(\beta)}$

$$\theta(\mathcal{H}) = \frac{1}{\chi}\int_{\mathcal{H}}^{\infty} e^{-\beta E}\, dE = \frac{1}{\beta\chi} e^{-\beta \mathcal{H}}$$

Then: $\int \theta(\mathcal{H})\, dV = \dfrac{1}{\beta\chi}\int e^{-\beta \mathcal{H}} dV = \dfrac{1}{\beta} = kT$

so we have the same result for the equipartition theorem in the CDF as we had in the MCDF.

Although we used classical notation here in our development of thermodynamics based on the CDF, we recall that we were able to give correspondences between classical SM and QSM in the case of the MCDF. However, in quantum mechanical problems we almost always use the CDF and most problems resolve to the calculation of the quantum mechanical partition function Q. The correspondence relations become:

$$\rho(\mathcal{H}) = \frac{e^{-\beta \mathcal{H}}}{Q(\beta, a)} \quad ; \quad Q = \text{Trace}\, e^{-\beta \mathcal{H}}$$

$$U = \bar{E} = \overline{\langle \mathcal{H} \rangle} = -\frac{\partial}{\partial \beta} \log Q$$

$$\overline{\langle F_a \rangle} = -\overline{\left\langle \frac{\partial H}{\partial a} \right\rangle} = \frac{1}{\beta Q}\frac{\partial Q}{\partial a} = \frac{1}{\beta}\frac{\partial}{\partial a} \log Q$$

and so on, just substituting Q for χ in all the classical expressions. However, one still must consider the correspondence limit between χ and Q.

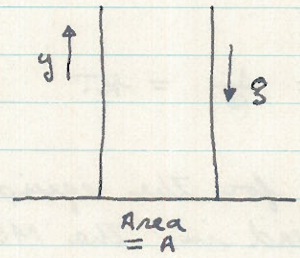
The equipartition theorem is completely absent in QM as there is no longer any meaning to p.s. and hence no corresponding operations can be done. This absence of equipartition lead to the development of QM through the experimental anomalies in C_v (Planck, Einstein, et al) different from those predicted by the equipartition theorem.

Let us now consider an example in classical SM.

Gas Molecules in Constant Gravitational Field

Consider a column of gas molecules in a constant gravitational potential infinitely high so no particles escape.

g = acceleration due to gravity



$$H = \sum_j \frac{p_j^2}{2m} + mg \sum_j y_j$$

The parameter "a" in the Hamiltonian is in this case g.

Since we have the parameter g we can, in an abstract sense, talk about a parameter-related force $\mathcal{F}_g$:

$$\mathcal{F}_g = -\frac{\partial H}{\partial g} = -m \sum_j y_j$$

Note that the centroid of the distribution of the gas in the column is given by:

$$h = \frac{1}{N} \sum_j y_j \quad \text{and thus:} \quad \mathcal{F}_g = -mNh$$

What is the single particle partition function ξ?

$$\xi(\beta, g) = \left(\frac{2\pi m}{\beta}\right)^{3/2} A \int_0^\infty e^{-\beta m g y}\, dy = \left(\frac{2\pi m}{\beta}\right)^{3/2} \frac{A}{\beta m g}$$

$$\log X = N \left\{ \log (2\pi m)^{3/2} \frac{A}{m} - \frac{5}{2} \log \beta - \log g \right\}$$

$$U = -\frac{\partial}{\partial \beta} \log X = \frac{5N}{2\beta} = \frac{5}{2} NkT$$

This may seem peculiar because we know that for a free particle gas, $C = \frac{3}{2} Nk$. However, here we have:

$$C_g = \frac{5}{2} Nk = \frac{3}{2} Nk + Nk$$

The extra Nk is needed because when the temperature is raised, work is done against gravity.

To show that this interpretation is correct, use the Equipartition Theorem:

$$\overline{y_g \frac{\partial H}{\partial y_g}} = \overline{mg\, y_g} = kT$$

note that as a result of the above:

$$\bar{h} = \frac{kT}{mg}$$

Then: $\overline{F}_g = -\frac{1}{\beta}\frac{N}{g} = -kT\frac{N}{g}$

obtained from $\overline{F}_g = \frac{1}{\beta}\frac{\partial}{\partial g}\log \chi$. But: $\overline{F}_g = -mN\bar{h}$ so that we get the same result as deduced from the equipartition theorem:

$$\boxed{\bar{h} = \frac{kT}{mg}}$$

which is a well known relation. The "barometer formula" can be deduced from the CDF directly:

$\rho = \frac{1}{\chi} e^{-\beta\sum\frac{p^2}{2m} - \beta mg\sum y}$ and we see that the atmospheric density goes as $e^{-\beta mgy}$.

The question of fluctuations in the generalized force $\overline{F}_a$ can be very important. We consider this now with later applications to the free gas and gas in a gravitational potential. We have:

$$\overline{F}_a = \frac{1}{\beta\chi}\frac{\partial\chi}{\partial a} = \frac{1}{\beta}\frac{\partial}{\partial a}\log\chi = \frac{-1}{\chi}\int\frac{\partial H}{\partial a}e^{-\beta H}dV$$

$$\overline{F_a^2} = \frac{1}{\chi}\int\left(\frac{\partial H}{\partial a}\right)^2 e^{-\beta H}dV$$

Notice that $\frac{\partial^2}{\partial a^2}\chi = \frac{\partial^2}{\partial a^2}\int e^{-\beta H}dV = \frac{\partial}{\partial a}\int\left(-\beta\frac{\partial H}{\partial a}\right)e^{-\beta H}dV$

$$= \int\left\{\beta^2\left(\frac{\partial H}{\partial a}\right)^2 - \beta\frac{\partial^2 H}{\partial a^2}\right\}e^{-\beta H}dV$$

Then:

$$\overline{F_a^2} = \frac{1}{\beta^2\chi}\frac{\partial^2}{\partial a^2}\chi + \frac{1}{\beta\chi}\int\frac{\partial^2 H}{\partial a^2}e^{-\beta H}dV$$

Then: $$\overline{\mathcal{F}_a^2} - \overline{\mathcal{F}_a}^2 = \frac{1}{\beta^2}\left\{\frac{1}{\chi}\frac{\partial^2}{\partial a^2}\chi - \left(\frac{1}{\chi}\frac{\partial \chi}{\partial a}\right)^2\right\} + \frac{1}{\beta\chi}\int \frac{\partial^2 H}{\partial a^2}e^{-\beta H}dV$$

Finally:
$$\boxed{\begin{aligned}\overline{\mathcal{F}_a^2} - \overline{\mathcal{F}_a}^2 &= \frac{1}{\beta^2}\frac{\partial^2}{\partial a^2}\log \chi + \frac{1}{\beta\chi}\int \frac{\partial^2 H}{\partial a^2}e^{-\beta H}dV \\ &= \frac{1}{\beta}\left\{\frac{\partial}{\partial a}\overline{\mathcal{F}_a} - \overline{\frac{\partial \mathcal{F}_a}{\partial a}}\right\}\end{aligned}}$$

because $\mathcal{F}_a = -\frac{\partial H}{\partial a}$ and $\overline{\frac{\partial \mathcal{F}_a}{\partial a}} = -\overline{\frac{\partial^2 H}{\partial a^2}}$

What are the fluctuations in the pressure of a gas?

$$p = \overline{\mathcal{F}_v}$$

$$\overline{\mathcal{F}_v^2} - p^2 = \overline{(\mathcal{F}_v - p)^2} = kT\left\{\frac{\partial p}{\partial v} - \overline{\frac{\partial \mathcal{F}_v}{\partial v}}\right\}$$

Note that $\overline{\frac{\partial \mathcal{F}_v}{\partial p}}$ cannot be obtained from thermodynamics alone. The perfect gas yields an infinite fluctuation because the collisions of the molecules with the container walls are δ functions.

Now let us look at our gravitational potential model:

$$\mathcal{F}_g = -Nmh \; ; \quad \overline{\mathcal{F}_g} = -Nm\frac{kT}{mg} = -\frac{NkT}{g}$$

since the Hamiltonian is linear in g we have $\frac{\partial^2 H}{\partial g^2} = 0$ and:

$$\overline{(\mathcal{F}_g - \overline{\mathcal{F}_g})^2} = kT\frac{\partial}{\partial g}\left(-\frac{NkT}{g}\right) = N\left(\frac{kT}{g}\right)^2$$

The relative fluctuation is:

$$\frac{\sqrt{\overline{(\mathcal{F}_g - \overline{\mathcal{F}_g})^2}}}{|\overline{\mathcal{F}_g}|} = \frac{\sqrt{N}\,\frac{kT}{g}}{N\,\frac{kT}{g}} = \frac{1}{\sqrt{N}}$$

so that the relative fluctuation in the hieght of the centroid goes as $\frac{1}{\sqrt{N}}$ like most fluctuations.

Problem ⑪

Consider a perfect gas of linear molecules possessing a dipole moment d:

$$H = \sum_j \left\{ \frac{p_j^2}{2m} + \frac{1}{2A}\left(p_{\theta_j}^2 + \frac{p_{\varphi_j}^2}{\sin^2\theta_j} \right) - \mathcal{E} d \cos\theta_j \right\}$$

$\mathcal{F}_\mathcal{E} = -\dfrac{\partial H}{\partial \mathcal{E}} = d \sum_j \cos\theta_j$ which represents the total dipole moment in the z direction.

Find: $U,\ C_\mathcal{E},\ \overline{\mathcal{F}_\mathcal{E}},\ \overline{\mathcal{F}_\mathcal{E}^2} - \overline{\mathcal{F}_\mathcal{E}}^2$

Note explicitly the values for $\mathcal{E} = 0$ and $\mathcal{E} = \infty$ and interpret.

Problem ⑫

Consider a collection of fixed particles of spin ½, neglecting any rotational kinetic energy.

$$H = -\sum_j \mathcal{E} d \, \sigma_{z j}$$

This Hamiltonian disobeys parity conservation so it is not right. Find the same things as in ⑪ except for the rotational and translational parts. Make a comparison with the appropriate parts of ⑪ and look at $\mathcal{E} \to 0$ and $\mathcal{E} \to \infty$ and interpret.

Problem ⑬

Consider a 1-D harmonic oscillator at temperature T. For each oscillator in the collection:

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2$$

Find and compare the classical and quantum values of $U,\ C,\ \overline{q^2}$ as functions of temperature. How does C compare with the prediction of the classical equipartition theorem?

As an optional problem, try to find the distribution function of q at temperature T for the classical and quantum cases.

Most General Rotational Partition Function

If we define the 3 Euler angles, α, β, γ, and the principle momenta of inertia A, B, C, we have for the Hamiltonian:

$$H = \frac{1}{2}\left\{ \frac{1}{A \sin^2\beta}\left[(p_\alpha - p_\gamma \cos\beta)\cos\gamma - p_\beta \sin\gamma \sin\beta \right]^2 \right.$$

$$\left. + \frac{1}{B \sin^2\beta}\left[(p_\alpha - p_\gamma \cos\beta)\sin\gamma + p_\beta \cos\gamma \sin\beta \right]^2 + \frac{p_\gamma^2}{C} \right\}$$

$$\chi = \int e^{-\beta H}\, dp_\alpha\, dp_\beta\, dp_\gamma\, d\alpha\, d\beta\, d\gamma$$

Make the Transformation: $p_\alpha' = \dfrac{p_\alpha - p_\gamma \cos\beta}{\sin\beta}$; $dp_\alpha = \sin\beta\, dp_\alpha'$

and then the orthogonal transformation:

$$\left.\begin{aligned} p_A &= p_\alpha' \cos\gamma - p_\beta \sin\gamma \\ p_B &= p_\alpha' \sin\gamma + p_\beta \cos\gamma \end{aligned}\right\} \quad \frac{\partial(p_\alpha'\, p_\beta)}{\partial(p_A\, p_B)} = 1$$

Then:

$$\chi = \int e^{-\frac{\beta}{2}\left\{ \frac{p_A^2}{A} + \frac{p_B^2}{B} + \frac{p_\gamma^2}{C} \right\}}\, dp_A\, dp_B\, dp_\gamma \sin\beta\, d\alpha\, d\beta\, d\gamma$$

The ranges of integration are:

$0 \leq \beta \leq \pi$
$0 \leq \alpha \leq 2\pi$
$0 \leq \gamma \leq 2\pi$

Then:

$$\chi = \left(\frac{2\pi^2 ABC}{\beta^3}\right)^{1/2} \int \sin\beta\, d\alpha\, d\beta\, d\gamma$$

or:

$$\boxed{\chi = 8\pi^2 (2\pi kT)^{3/2} \sqrt{ABC}}$$

LECTURE 25 : 1-9-62

Problem ⑭

Consider the Linear Quantum Mechanical Rotator whose eigenvalues are:

$$E_J = \frac{\hbar^2}{2A} J(J+1) \quad ; \quad J = 0, 1, 2, \cdots$$

Each level is $2J+1$ - fold degenerate. Find: $Q(\beta)$, U, and C for N rotators. The sum over the eigenvalues cannot be carried out exactly so that we will have to consider high and low temperature limits, that is, as $\frac{\hbar^2}{2AkT} \to 0$ and ∞. Compare with classical results.

Most SM problems become problems in summation or integration and are consequently usually either trivial or impossible. The trivial problems mostly involve independent systems or separable Hamiltonians. Problems with interaction potentials are very difficult to handle which is one of the stumbling blocks in the theory of phase changes in liquids.

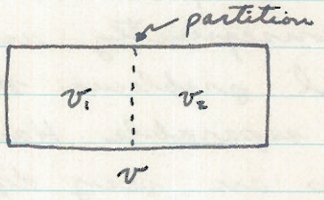
Up to now we have discussed systems whose number of degrees of freedom remain fixed. However, many chemical systems do not behave like this, for example, the hydrogen reaction $H + H \to H_2$ reduces the number of degrees of freedom from 6 to 5 because H_2 behaves as a rigid rotator.

Another reaction that does not preserve the number of degrees of freedom is that of pair production which requires energies of the order of 1 MEV. It is possible to have stellar temperatures like this energy and in such cases we have dynamic thermal equilibrium between positrons, electrons and photons with the number of degrees of freedom continually changing.

If we consider systems of photons alone, for example, two isolated bodies, the number of photons in either one are not conserved because of radiation, and hence the degrees of freedom are not conserved.

We will treat the problem of the non-conservation of the degrees of freedom by analogy to the canonical distribution in that we will say that we will not specify the number of particles exactly. That is, we will consider a very large system so that the number of particles is fixed, but we will allow the free exchange of particles among the subsystems. In this way, we will get an intensive parameter, the same for all subsystems (like β), which is called the chemical potential.

We give a discussion of this problem due to Gibbs. Consider a volume of gas, v, in which the total number of particles is specified which is contained in a box. We let the gas come to equilibrium and then slide in a partition.



System in equilibrium

We have:

$$S_1 = k \log V_1(E_1) \approx k \log \Omega_1(E_1) \approx k\left(\beta E_1 + \log \chi_1(\beta v_1)\right)$$

Using the well-known relations for the free gas:

$$\frac{S_1}{k} = \beta \frac{3}{2} N_1 kT + N_1 \log (2\pi m kT)^{3/2} v_1$$

$$= N_1 \left\{ \frac{3}{2} + \frac{3}{2} \log 2\pi m kT + \log v_1 \right\}$$

We have a similar expression for $\frac{S_2}{k}$

We now pull out the partition, then $v = v_1 + v_2$, $N = N_1 + N_2$ and the entropy is given by:

$$\frac{S}{k} = (N_1 + N_2) \left\{ \frac{3}{2} + \frac{3}{2} \log 2\pi m kT + \log (v_1 + v_2) \right\}$$

We can now compare S with $S_1 + S_2$.

We get:

$$\frac{S-S_1-S_2}{k} = (N_1+N_2)\log(v_1+v_2) - N_1\log v_1 - N_2\log v_2$$

Since the density is uniform and does not change, we may write:

$$\frac{N_1}{v_1} = \frac{N_2}{v_2} = \frac{N_1+N_2}{v_1+v_2} = n$$

Then:

$$\frac{S-S_1-S_2}{k} = (N_1+N_2)\log(N_1+N_2) - N_1\log N_1 - N_2\log N_2$$

$$- \underbrace{(N_1+N_2-N_1-N_2)}_{0}\log n$$

$$= N_1\log\frac{N_1+N_2}{N_1} + N_2\log\frac{N_1+N_2}{N_2} > 0$$

because $N_1+N_2 > N_1, N_2$. We thus see that the entropies are <u>not additive</u>. Now we have said that S is additive for two dynamically independent systems. However, here we apparently have interaction between two systems of identical particles. What makes this more strange is the fact that the partition is really imaginary.

Gibbs found this paradox and noted that previously no one had considered changes in the degrees of freedom of the system. Thus he concluded that the entropy must include an additive constant that depends on the degrees of freedom, that is:

$$\frac{S}{k} \to \frac{S}{k} + C(N)$$

the condition on $C(N)$ being such that it make the entropy additive:

$$(N_1+N_2)\log(N_1+N_2) - N_1\log N_1 - N_2\log N_2$$

$$= C(N_1+N_2) - C(N_1) - C(N_2)$$

This is a difference equation whose most general solution is:

$$C(N) = N\log N - JN$$

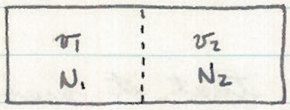
suppose we choose $I = 1$, then $C(N) = N \log N - N \approx \log N!$. By this route, Gibbs concluded that it would be a good idea to include $N!$ in the general SM definition of the entropy:

$$S = k \log \frac{V(E)}{N!} \approx k \log \frac{\Omega(E)}{N!} \approx k \left\{ \beta U + \log \frac{\chi(\beta)}{N!} \right\}$$

$N!$ downstairs so it subtracts off the right amount so that S adds. We see that $N!$ makes no difference in the usual thermodynamic relations unless the degrees of freedom are changing. The presence of $N!$ is here only suggested and not proved as we have used the crutch of the free gas. However, Gibbs took it to be general.

Gibbs noted that if we permuted the molecules among fixed coordinates in p.s. (or a <u>specific phase</u>) we always come out with the same result on the laboratory scale. Gibbs said that the $N!$ permutations correspond to the same <u>generic phase</u>. Thus it would not be surprising that if when integrating over p.s. one must divide by $N!$ to get rid of the corresponding configurations. It is at this point that some people claim that Gibbs anticipated quantum mechanics.

We will now show more formally and completely the way $N!$ arises. Consider a complete system, in equilibrium, subdivided into two parts as before only now we need know nothing of nature of interaction among the particles.



Label each particle by γ such that $1 \leq \gamma \leq N_1$ in the first part and $N_1 + 1 \leq \gamma \leq N_1 + N_2 = N$ in the second part.
Define:

$\omega_1(\epsilon_\gamma)$ = structure function for γ^{th} molecule in v_1 with energy ϵ^γ and likewise for v_2.

$$\omega(\epsilon_\gamma) = \omega_1(\epsilon_\gamma) + \omega_2(\epsilon_\gamma)$$

The total for each of the separate systems is:

$$\Omega_1(E_1) = \int \delta(E_1 - \sum_{\gamma=1}^{N_1} \epsilon_\gamma) \prod \omega_1(\epsilon_\gamma) \prod d\epsilon_\gamma$$

$$\Omega_2(E_2) = \int \delta(E_2 - \sum_{\gamma=N_1+1}^{N_1+N_2} \epsilon_\gamma) \prod \omega_2(\epsilon_\gamma) \prod d\epsilon_\gamma$$

while for the composite system it is.

$$\Omega(E) = \int \delta(E - \sum_{\gamma=1}^{N_1+N_2} \epsilon_\gamma) \prod_{\gamma=1}^{N_1+N_2=N} \{\omega_1(\epsilon_\gamma) + \omega_2(\epsilon_\gamma)\} \prod d\epsilon_\gamma$$

We expand the product and pick out ω_1 and ω_2 products in the sum. Pick out the independent products with the binomial coefficient. Of course now N_1 and N_2 are not the same as before so call them N_1' and N_2'. Then:

$$\Omega(E) = \sum_{N_1'=0}^{N} \int \delta(E - \sum \epsilon(\gamma)) \frac{N!}{N_1'!\,(N-N_1')!} \prod_{\gamma=1}^{N_1'} \omega_1(\epsilon_\gamma) \prod_{\gamma=N_1'+1}^{N} \omega_2(\epsilon_\gamma) \prod d\epsilon_\gamma$$

$$= \sum_{N_1'=0}^{N} \int \delta(E - E_1' - E_2')\, dE_1'\, dE_2'\ \delta(E_1' - \sum_{\gamma=1}^{N_1'} \epsilon_\gamma) \prod \omega_1(\epsilon_\gamma)\ \delta(E_2' - \sum_{\gamma=N_1'+1}^{N} \epsilon_\gamma)$$

$$\cdot \prod \omega_2(\epsilon_\gamma) \prod d\epsilon_\gamma \cdot \frac{N!}{N_1'!\,(N-N_1')!}$$

Then:

$$\Omega(E) = \sum_{N_1'=0}^{N} \int \delta(E - E_1' - E_2')\, \Omega_1(N_1' E_1')\, \Omega_2(N-N_1', E_2')\, dE_1'\, dE_2'$$

$$\cdot \frac{N!}{N_1'!\,(N-N_1')!}$$

Or:

$$\frac{\Omega(E)}{N!} = \sum_{N_1'=0}^{N} \int \frac{\Omega_1(N_1' E_1')}{N_1!} \quad \frac{\Omega_2(N-N_1', E-E_1')}{(N-N_1')!}\, dE_1$$

which is an exact result for the situation when the partition is removed. In practice we find the integrand highly peaked at some value of E_1 and the summand peaked for some value of N_1. We will sketch how this comes about. Call W the integrand-summand and take the logarithm:

$$\log W = \log\left[\frac{\Omega_1(E_1 N_1)}{N_1!}\;\frac{\Omega_2(E-E_1\; N-N_1)}{(N-N_1)!}\right] \;;\; N_2 = N-N_1 \;,\; E_2 = E-E_1$$

$$\delta \log W = \left[\frac{\partial}{\partial E_1}\log\frac{\Omega_1(E_1 N_1)}{N_1!} - \frac{\partial}{\partial E_2}\log\frac{\Omega_2(E_2 N_2)}{N_2!}\right]dE_1$$

$$+\left[\frac{\partial}{\partial N_1}\log\frac{\Omega_1(E_1 N_1)}{N_1!} - \frac{\partial}{\partial N_2}\log\frac{\Omega_2(E_2 N_2)}{N_2!}\right]dN_1$$

For a peak to occur, the coefficients of dE_1, dN_1 must vanish:

$$\underbrace{\frac{\partial}{\partial E_1}\log\Omega_1(E_1 N_1)}_{\frac{1}{kT_1}} = \underbrace{\frac{\partial}{\partial E_2}\log\Omega_2(E_2 N_2)}_{\frac{1}{kT_2}}$$

; therefore $T_1 = T_2$ which is known as the first condition.

Consider the coefficient of dN_1 on the perfect gas model where:

$$\log\frac{\Omega_1}{N_1!} = \log\frac{2m(2\pi m E_1)^{\frac{3N_1}{2}}\, v_1^{N_1}}{\Gamma\left(\frac{3N_1}{2}\right)N_1!}$$

$$\frac{\partial}{\partial N_1}\log\frac{\Omega_1}{N_1!} = \frac{3}{2}\log\left[2\pi m kT\right] + \log\frac{v_1}{N_1}$$

so when we take the other coefficient we have the second condition that $\frac{v_1}{N_1} = \frac{v_2}{N_2}$ or that the densities are equal.

We now make an approximation, that of neglecting the widths of the peaks, and which is appropriate for thermodynamics:

$$\frac{\Omega(E)}{N!} \doteq \frac{\Omega_1(N_1E_1)}{N_1!}\;\frac{\Omega_2(N_2E_2)}{N_2!}$$

times a correction factor which only adds a small number when the logarithm is taken.

Hence we evaluate the integral and sum by taking the integrand-summand at the equilibrium values of E_1 and N_1. Then:

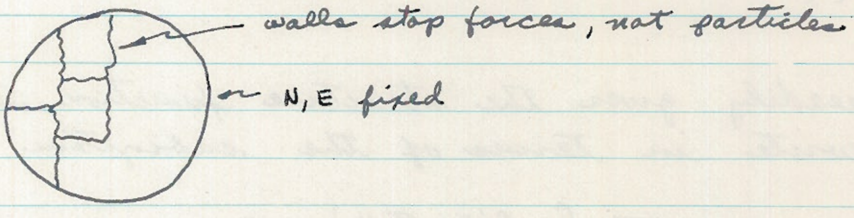
$$\log\frac{\Omega(NE)}{N!} = \log\frac{\Omega_1(N_1E_1)}{N_1!} + \log\frac{\Omega_2(N_2E_2)}{N_2!}$$

which is the same result as Gibbs, however, we still have used the perfect gas as a crutch, next we will generalize and introduce the grand canonical ensemble.

LECTURE 26 : 1-11-62

The Grand Canonical Ensemble

We consider a supersystem composed of subsystems. Particles may pass from subsystem to subsystem, but not forces. The total energy and number of particles of the supersystem is fixed.



Saying that no forces pass through boundary is like neglecting surface effects or intermolecular effects across the boundary.

Consider the whole system whose phase point is x.

$$\rho_{E,N} = \frac{\delta(E - H(x))}{\Omega(E,N)}$$

Consider subsystem j with N_j particles and assume there are n subsystems. Then we have a set of numbers $\{N_j\}$. For each subsystem, construct a p.s. of $6N_j$ dimensions and since each subsystem can be described by a point x_j we have the set of points $\{x_j\}$. Now we are able to permute the particles around among the various subsystems as long as we are willing to continuously relable the particles. Since we don't consider permutations among the particles of the subsystem itself as representing anything new, we have for the number of ways of arrangement:

$$\frac{N!}{N_1!\,N_2!\,N_3!\cdots} = \frac{N!}{\prod N_j!}$$

This is the number of points in x corresponding to one set $\{x_j\}$. We will now define the <u>generic density</u>:

$$\rho(\{N_j\},\{x_j\}) = \underbrace{\frac{N!}{\Pi N_j!}}_{\text{normalization}} \; \frac{\delta(E - \Sigma H_j(x_j))}{\Omega(E,N)} \; \underbrace{\delta_{N,\Sigma N_j}}_{\text{assures that the sum total is correct}}$$

The normalization condition is:

$$\sum_{\{N_j\}} \int \rho(\{N_j\},\{x_j\}) \prod_{j=1}^{n} dV_j = 1$$

This readily gives the structure function $\Omega(E,N)$ which we then can write in terms of the subsystem structure functions.

$$\Omega(E,N) = N! \sum_{\{N_j\}} \int \frac{\delta(E-\Sigma H_j)}{\Pi N_j!} \delta_{N,\Sigma N_j} \, \Pi \, dV_j$$

$$= N! \sum_{\{N_j\}} \delta_{N,\Sigma N_j} \int \delta(E - \Sigma E_j) \, \Pi \underbrace{\frac{\delta(E_j - H_j)}{N_j!}}_{\substack{\text{subsystem} \\ \text{structure} \\ \text{functions}}} \Pi \, dV_j \, \Pi \, dE_j$$

Then:

$$\boxed{\Omega(E,N) = N! \sum_{\{N_j\}} \delta_{N,\Sigma N_j} \int \delta(E - \Sigma E_j) \, \Pi \frac{\Omega_j(E_j N_j)}{N_j!} \, \Pi \, dE_j}$$

This is the composition law.

Now we would like to find a density function for each subsystem. We reduce ρ above to subsystem 1 by integrating and summing over all but 1. We get:

$$P(N_1, x_1) = \frac{N!}{N_1! \, \Omega(EN)} \sum_{\substack{\{N_j\} \\ j=2\to n}} \delta_{N-N_1, \sum_{j=2}^{n} N_j} \frac{\int \delta(E - H_1 - \sum_{j=2}^{n} H_j) \prod_{j=2}^{n} dV_j}{\prod_{j=2}^{n} N_j!}$$

$$= \frac{1}{N_1!} \frac{N!}{\Omega(EN)} \frac{\Omega^{(N-1)}(E - H_1, N - N_1)}{(N - N_1)!}$$

$\Omega^{(N-1)}$ is the complementary structure function. The above is very analogous to some old results.

We now would like to produce an asymptotic form for the total structure function or more exactly $\frac{\Omega(E,N)}{N!}$. We recall the ordinary partition function:

$$\chi(\alpha, N) = \int e^{-\alpha H} dV = \int e^{-\alpha E} \Omega(E,N)\, dE$$

We define a generating function:

$$\boxed{\Xi(\alpha, z) = \sum_{N=0}^{\infty} z^N \frac{\chi(\alpha,N)}{N!}}$$

If we know Ξ, we can find $\chi(\alpha, N)$ via Cauchy's Integral Formula. The advantage of the generating function is that we need not know precisely the value of N as it is "summed out" just as E is "integrated out" by the conventional definition of χ. Therefore, we see that Ξ behaves like a partition function and indeed it is called the <u>Grand Partition Function.</u> For mathematical consistency we must define the physically meaningless $\chi(\alpha, 0)$ to be 1. In terms of the structure function:

$$\boxed{\Xi(\alpha, z) = \sum_{N=0}^{\infty} \frac{z^N}{N!} \int_0^{\infty} e^{-\alpha E} \Omega(E,N)\, dE}$$

The role of Ξ as a partition function is now clear. We will later find the physical meaning of z as we did earlier for α. Now, written in terms of the subsystems:

$$\Xi(\alpha, z) = \sum_{N=0}^{\infty} \frac{z^N}{N!} \int e^{-\alpha E} N! \sum_{\{N_j\}} \delta_{N, \Sigma N_j} \int \delta(E - \Sigma E_j) \prod \frac{\Omega_j(E_j N_j)}{N_j!} \prod dE_j\, dE$$

We will have in the expression $\prod_j \frac{e^{-\alpha E_j} \Omega_j(E_j N_j)}{N_j!}$ because of the Dirac δ function and because the Kronecker δ makes $N \to \Sigma N_j$ we also get a $z^{\Sigma N_j}$; hence:

$$\Xi(\alpha, z) = \sum_{\{N_j\}} \int \prod_j z^{N_j} \frac{e^{-\alpha E_j} \Omega_j(E_j N_j)}{N_j!} \prod dE_j$$

$$= \prod_{j=1}^{n} \sum_{N_j=0}^{\infty} \frac{z^{N_j}}{N_j!} \int e^{-\alpha E_j} \Omega_j(E_j N_j)\, dE_j = \prod_{j=1}^{n} \Xi_j(\alpha, z)$$

Hence the composition law still holds for the grand partition function.

Now, to get the asymptotic structure function we proceed as follows:

$$\frac{\Omega(E,N)}{N!} = \sum_{\{N_j\}} \delta_{N,\Sigma N_j} \int \delta(E - \Sigma E_j) \prod \frac{\Omega_j(E_j N_j)}{N_j!} \prod dE_j$$

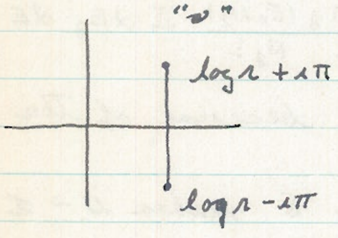
$$= \frac{1}{2\pi i} \sum_{\{N_j\}} \delta_{N,\Sigma N_j} \int e^{\alpha E} \prod \frac{\chi_j(\alpha)}{N_j!} \, d\alpha$$

using the integral representation of the Dirac δ function: $\delta(x) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{\alpha x} d\alpha$. We now use the integral representation for the Kronecker δ function: $\delta_{N,N'} = \frac{1}{2\pi i} \oint \frac{dz}{z^{N-N'+1}}$ where $N' = \Sigma N_j$. Then:

$$\frac{\Omega}{N!} = \frac{1}{(2\pi i)^2} \oint\int \frac{e^{\alpha E}}{z^{N+1}} \sum_{\{N_j\}} \prod_j \frac{z^{N_j} \chi_j(\alpha)}{N_j!} \, d\alpha \, dz$$

$$= \frac{1}{(2\pi i)^2} \oint\int \frac{e^{\alpha E}}{z^{N+1}} \prod_j \Xi_j(\alpha z) \, d\alpha \, dz$$

We must now find a saddle point in two variables simultaneously. Define for a new independent variable $z = e^{\nu}$, $dz = z\,d\nu$. This straightens out the contour in z from a circle to a straight line in the ν plane.



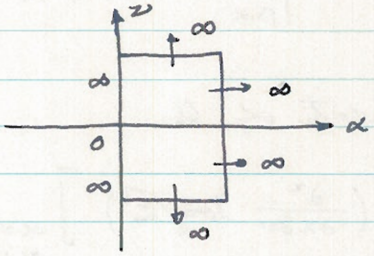
$$\frac{\Omega(EN)}{N!} = \frac{1}{(2\pi i)^2} \iint e^{\alpha E - \nu N + \log \Xi(\alpha,\nu)} \, d\alpha \, d\nu$$

If we consider α and ν real variables, we can show that there is one (or at least one) minimum in the $\alpha - \nu$ plane.

Consider the exponent as a function of α and ν:

$$g(\alpha,\nu) = \alpha E - \nu N + \log \Xi(\alpha,\nu)$$

We see that this has the following behavior in the $\alpha-\nu$ plane:



hence a minimum must exist

The minima of $g(\alpha,\nu)$ are given by the conditions:

$$E = -\frac{\partial}{\partial\alpha}\log\Xi(\alpha\nu) \quad ; \text{ roots } \alpha = \beta$$

$$N = \frac{\partial}{\partial\nu}\log\Xi(\alpha\nu) \quad ; \text{ roots } \nu = \beta\mu$$

Also, we have to show that the second variation in $g(\alpha,\nu)$ is positive definite at $\alpha=\beta$ and $\nu=\beta\mu$ in order to prove that we are at a minimum. We expand around β, $\beta\mu$ and examine the quadratic term:

$$Q = \left\{(\alpha-\beta)^2\frac{\partial^2}{\partial\alpha^2} + 2(\alpha-\beta)(\nu-\beta\mu)\frac{\partial^2}{\partial\alpha\,\partial\nu} + (\nu-\beta\mu)^2\frac{\partial^2}{\partial\nu^2}\right\}\log\Xi \geq 0$$

Consider a likely distribution function (multivariate):

$$\frac{e^{-\alpha E+\nu N}\,\Omega(EN)}{N!\;\Xi(\alpha,\nu)}$$

We claim that (in analogy to what we know already):

$$\frac{\partial^2}{\partial\alpha^2}\log\Xi = \overline{(E-\bar{E})^2} \quad ; \quad \frac{\partial^2}{\partial\nu^2}\log\Xi = \overline{(N-\bar{N})^2}$$

$$\frac{\partial^2}{\partial\alpha\,\partial\nu}\log\Xi = -\overline{(E-\bar{E})(N-\bar{N})} \quad (\text{cross-correlation})$$

Then:

$$Q = \overline{\left\{(\alpha-\beta)(E-\bar{E}) - (\nu-\beta\mu)(N-\bar{N})\right\}^2} \geq 0, \text{ obviously.}$$

Hence we are dealing with a simultaneous saddle point in the complex variables α and ν. We now evaluate using the usual saddle point method and get:

$$\frac{\Omega(EN)}{N!} = e^{\beta E - \beta\mu N} \, \Xi(\alpha,\nu)\Big|_{\substack{\beta \\ \beta\mu}} \; \frac{1}{2\pi\sqrt{\Delta}}$$

where Δ is the discriminant of Q:

$$\Delta = \left[\frac{\partial^2}{\partial\alpha^2}\log\Xi \, \frac{\partial^2}{\partial\nu^2}\log\Xi - \left(\frac{\partial^2}{\partial\alpha\,\partial\nu}\log\Xi\right)^2 \right]_{\substack{\alpha=\beta \\ \nu=\beta\mu}}$$

The Δ term may be dropped when taking logarithms. We now may write for the first subsystem:

$$\boxed{P(N_1 X_1) = \frac{e^{-\beta(H_1 - \mu N_1)}}{N_1! \, \Xi(\beta,\mu)}}$$

where $H_1 = H_1(N_1 X_1)$.

$P(N_1 X_1)$ is called the Grand Canonical Distribution Function. It got this name from Gibbs because it originally represented an ensemble of ensembles.

LECTURE 27 : 1-16-62

We now write down some thermodynamic laws relating to systems in which the number of particles is able to change. The first law is:

$$\boxed{d\mathcal{U} = \delta Q - \delta W + \mu\, d\mathcal{N}}$$

μ is the chemical potential and $\mathcal{N}$ is in the same sense as $\mathcal{U}$, that is, $\mathcal{N} = \bar{N}$. The second law and the work term are:

$$\boxed{\delta Q = T dS \quad ; \quad \delta W = p dV}$$

Then, if we take T, V, and $\mathcal{N}$ as independent variables, we can deal with the Helmholtz Free Energy:

$$F = \mathcal{U} - TS \ ; \qquad dF = d\mathcal{U} - d(TS) = TdS - pdV + \mu d\mathcal{N} - d(TS)$$

or: $$\boxed{dF = -SdT - pdV + \mu d\mathcal{N}}$$

This means we can immediately identify:

$$\boxed{S = -\frac{\partial F}{\partial T} \ ; \ p = -\frac{\partial F}{\partial V} \ ; \ \mu = \frac{\partial F}{\partial \mathcal{N}}}$$

We can find F from X or Ξ.

Suppose we choose as independent variables T, P, and $\mathcal{N}$. Then we use the Gibbs Free Energy:

$$\boxed{G = F + pV = \mathcal{U} - TS + pV}$$

$$\boxed{dG = dF + d(pV) = -SdT + Vdp + \mu d\mathcal{N}}$$

$$\boxed{S = -\frac{\partial G}{\partial T} \ ; \ V = \frac{\partial G}{\partial p} \ ; \ \mu = \frac{\partial G}{\partial \mathcal{N}}}$$

A virtue of G is that two of the independent variables are intensive (T, p) and one extensive ($\mathcal{N}$). This means that if we have a homogeneous medium characterized by an amount of stuff $\mathcal{N}$, we should be able to write:

$$G(T, p, \lambda\mathcal{N}) = \lambda\, G(T, p, \mathcal{N})$$

which we expect to hold as long as boundary effects can be neglected. Suppose we differentiate with respect to λ and then let $\lambda \to 1$. We get, for a homogeneous medium only:

$$\mathcal{N}\frac{\partial G}{\partial \mathcal{N}} = G$$

Then: $G = \mu\mathcal{N}$ or $\mu = \frac{G}{\mathcal{N}} = \frac{1}{\mathcal{N}}(U - TS + pV)$

How many independent variables does a homogeneous have? It can have two intensive but not three because $\mu = \mu(p, T)$ as seen above so μ does not form an independent intensive quantity from p and T.

$$dG = d(\mu\mathcal{N}) = -S\,dT + V\,dp + \mu\, d\mathcal{N}$$

Hence: $\mathcal{N}\,d\mu = -S\,dT + V\,dp$

or: $d\mu = -\frac{S}{\mathcal{N}}\,dT + \frac{V}{\mathcal{N}}\,dp$

If we define $v = V/\mathcal{N}$ as the specific volume, we have the relations:

$$\boxed{\left(\frac{\partial\mu}{\partial p}\right)_T = V/\mathcal{N} = v \quad ; \quad \left(\frac{\partial\mu}{\partial T}\right)_p = -\frac{S}{\mathcal{N}}}$$

for homogeneous mediums only.

There is another useful function like F and G but unlike F and G it has no name. Its independent variables are T, V, and μ. Define it as:

$$\boxed{J = F - \mu\mathcal{N}}$$

Then:

$$\boxed{dJ = dF - d(\mu\mathcal{N}) = -S\,dT - p\,dV - \mathcal{N}\,d\mu}$$

$$\boxed{S = -\frac{\partial J}{\partial T} \; ; \quad p = -\frac{\partial J}{\partial V} \quad ; \quad \mathcal{N} = -\frac{\partial J}{\partial \mu}}$$

For homogeneous media, we have the special relation:

$$\boxed{J = F - G = -pV}$$

This has a very elegant application as if we know J, we automatically know the equation of state.

We now connect the foregoing remarks about thermodynamics with our results in SM on the grand canonical ensemble. For a subsystem we define $\mathcal{N} = \bar{N}$ in the same manner that we have $\mathcal{U} = \bar{E}$, but now we describe the subsystem with the GCDF.

$$\boxed{\mathcal{N} = \bar{N} = \frac{1}{\beta}\frac{\partial}{\partial \mu}\log \Xi(\beta,\mu)}$$

now, however, when we try taking $\frac{\partial}{\partial\beta}\log\Xi$ we get:

$$\boxed{\bar{E} - \mu\bar{N} = \mathcal{U} - \mu\mathcal{N} = -\frac{\partial}{\partial\beta}\log\Xi(\beta\mu)}$$

These two equations determine β and μ. Also we have for an equation of state:

$$\boxed{p = \overline{\mathcal{F}_V} = -\overline{\frac{\partial H}{\partial V}} = \frac{1}{\beta}\frac{\partial}{\partial V}\log\Xi(\beta,\mu)}$$

We now develop the full correspondence with thermodynamics. Consider $\log\Xi$ to be a function of β, μ, and V and form: $d\log\Xi$:

$$d\log\Xi = -(U-\mu\mathcal{N})\,d\beta + \beta p\,dV + \beta\mathcal{N}\,d\mu$$

now form:

$$d\left\{\beta(U-\mu\mathcal{N}) + \log\Xi\right\} = \beta\left\{dU - d(\mu\mathcal{N}) + p\,dV + \mathcal{N}\,d\mu\right\}$$

$$= \beta\left\{dU + p\,dV - \mu\,d\mathcal{N}\right\}$$

which we recognize as being the First Law of Thermodynamics:

$$\frac{1}{\beta}\,d\left\{\beta(U-\mu\mathcal{N}) + \log\Xi\right\} = \delta Q = T\,dS$$

This once again suggests that $\beta = \frac{1}{kT}$ through its action as an integrating factor. Thus:

$$\boxed{\beta = \frac{1}{kT}\ ; \quad \frac{S}{k} = \beta(U-\mu\mathcal{N}) + \log\Xi}$$

Then from: $TS = U - \mu\mathcal{N} + kT\log\Xi$, we have:

$$\boxed{J = -kT\log\Xi(\beta,\mu)}$$

which is exactly analogous to $F = -kT\log X(\beta)$. If we know J, we can find S, p, and $\mathcal{N}$ via relations already derived.

We now consider fluctuations in $U-\mu\mathcal{N}$: In analogy to the previous considerations on the fluctuations in U, we may immediately write:

$$\boxed{\begin{aligned}
\overline{(E-\mu N-\bar{E}+\mu\bar{N})^2} &= \frac{\partial^2}{\partial\beta^2}\log\Xi = -\frac{\partial}{\partial\beta}(U-\mu\mathcal{N}) \\
\overline{(N-\bar{N})^2} &= \frac{1}{\beta^2}\frac{\partial^2}{\partial\mu^2}\log\Xi = \frac{1}{\beta}\frac{\partial}{\partial\mu}\mathcal{N} \\
\overline{(N-\mathcal{N})^2} &= \frac{1}{\beta}\left(\frac{\partial\mathcal{N}}{\partial\mu}\right)_{T,V}
\end{aligned}}$$

We can put $\overline{(N-n)^2}$ in a more convenient form. Consider:

$$\left(\frac{\partial v}{\partial \mu}\right)_T = \left(\frac{\partial}{\partial \mu}\,\frac{V}{n}\right)_{T,V} = -\frac{V}{n^2}\left(\frac{\partial n}{\partial \mu}\right)_{T,V} = \left(\frac{\partial v}{\partial p}\right)_T \left(\frac{\partial p}{\partial \mu}\right)_T$$

Recall : $\left(\frac{\partial p}{\partial \mu}\right)_T = \frac{n}{V}$ (formerly we had the reciprocal)

Then:

$$\left(\frac{\partial v}{\partial \mu}\right)_T = \frac{1}{n}\left(\frac{\partial V}{\partial p}\right)_{T,n} \frac{n}{V}$$

Finally we have:

$$\left(\frac{\partial n}{\partial \mu}\right)_{T,V} = -\left(\frac{n}{V}\right)^2 \left(\frac{\partial V}{\partial p}\right)_{n,T}$$

so we can write :

$$\overline{(N-n)^2} = -kT\left(\frac{n}{V}\right)^2\left(\frac{\partial V}{\partial p}\right)_{n,T}$$

or; because the bulk modulus of compressibility is:

$$\kappa = -\frac{1}{V}\frac{\partial V}{\partial p}$$

we can write :

$$\boxed{\frac{\overline{(N-n)^2}}{n^2} = \frac{kT}{V}\kappa}$$

This also shows that κ is intrinsically positive, something that cannot be done with thermodynamics alone.

The Grand Canonical Ensemble in Quantum Mechanics

We now consider those situations in QM where we will have to make use of the grand canonical ensemble.

It is in general not possible that all quantum states that arise mathematically can be found in nature. The ordering of a number of electrons according to the Pauli Exclusion Principle among the mathematically possible states is an example of this.

The Pauli Exclusion Principle allows only one particle to occupy a single state at one time. One can, in an ad hoc fashion, incorporate this principle into the QM formalism by the formation of antisymmetric wave functions for a collection of electrons composed of single electron wave functions. Particles that have antisymmetric wave functions are called fermions and obey Fermi-Dirac statistics (FD). If a collection of particles have symmetric wave functions are called bosons and obey Bose-Einstein statistics (BE or EB).

The Pauli exclusion principle cannot be derived on the basis of NRQM. There are arguments in RQM that particles of spin 1/2 are antisymmetric and those with integer spin are symmetric. In the quantum field theory, one can show that if the reverse were true, it would lead to mathematical inconsistencies.

Let us now consider explicitly the quantum mechanics of a non-interacting, many electron system:

$$H \psi(r_1 \cdots r_N) = E\, \psi(r_1 \cdots r_N)$$

Suppose we perform a permutation operation

$$P_{12}\, \psi(r_1 r_2 \cdots) = \psi(r_2 r_1 \cdots)$$

Since H is invariant under the operation of P, we can write:

$$H P_{12} \psi = E P_{12} \psi$$

Thus, if the state ψ is non-degenerate, we must have:

$$P_{12}\psi = C\psi \;;\quad P_{12}^2 \psi = C^2 \psi = \psi \;;\quad C^2 = 1 \text{ or } C = \pm 1$$

This indicates that symmetry or antisymmetry may be present but it does not prove it.

We now discuss how we can form symmetric or antisymmetric wave functions.

Antisymmetry: One way that we can form antisymmetric states is to write:

$$\langle r_1 \cdots r_N | \rangle = \psi(r_1 \cdots r_N) = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \, P \, \varphi_1(r_1) \cdots \varphi_N(r_N)$$

where: $(-1)^P = \begin{cases} 1 & \text{even permutations} \\ -1 & \text{odd permutations} \end{cases}$

and the φ's are single electron wave functions. Another way to generate antisymmetric wave functions is by determinants (states):

$$\langle r_1 \cdots r_N | \rangle = \frac{1}{\sqrt{N!}} \det \left[\varphi_i(r_j) \right]$$

which of course vanishes unless all φ's are different and hence this satisfies the PEP.

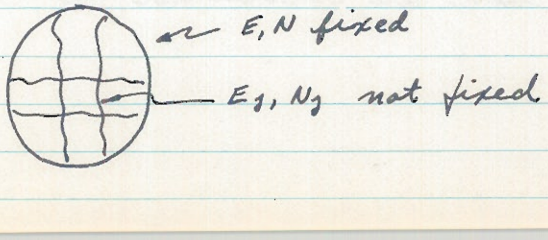
Symmetric: We may write these wave functions as:

$$\langle r_1 \cdots r_N | \rangle = \frac{1}{\sqrt{N! \, n_1! \, n_2! \cdots}} \sum_P P \, \varphi_1(r_1) \cdots \varphi_N(r_N)$$

where n_1 = number of repetitions of φ_1, etc. Using this type of wave function one cannot tell which particle is in which state.

For the antisymmetric case, the number of possible states is cut way down and we get nothing new on exchanging particles while for the symmetric case we get nothing new upon exchange but now no states are excluded.

We now try to apply the GCDF to the quantum statistics of bosons and fermions. Consider the collection of subsystems into a system as we did classically:



Each subsystem is specified by its state vector $|E_j, N_j\rangle$ while the supersystem is specified by $|E, N\rangle$. When we talk about all the subsystems, we have the set of state vectors $\{|E_j, N_j\rangle\}$. We see that only one state $|E, N\rangle$ can be constructed for the supersystem. Also, we must retain any symmetry or antisymmetry properties under interchanges of particles within and between subsystems.

We now invoke the same sort of arguments as used before when dealing with QSM. We use energy units of ϵ and take for the density operator:

$$\rho = \frac{\sum_k |E, N, k\rangle\langle E, N, k|}{W(E, N)}$$

Included in the $\sum_k$ are the sets $\{N_j\}$ and $\{E_j\}$ for the subsystems. We have for the total structure function, subject to the usual constraints:

$$W(E, N) = \sum_{\{N_j\}} \sum_{\{E_j\}} \delta_{N, \Sigma N_j} \, \delta_{E, \Sigma E_j} \prod_{j=1}^{n} W_j(E_j, N_j)$$

We also have, in analogy to previous results:

$$P^{(1)}_{N_1 E_1} = \frac{W^{(n-1)}(E - E_1, N - N_1)}{W^{(n)}(E, N)}$$

$$Q(\alpha, N) = \sum_E e^{-\alpha E} \, W(E, N)$$

Now, in the same spirit as the classical case, we define the grand partition function as:

$$\boxed{\Xi = \sum_N z^N Q(\alpha, N)}$$

Note that no factorials appear here as contrasted to the classical Ξ.

We now introduce integral representations for the Kronecker δ's:

$$\delta_{EE'} = \frac{\epsilon}{2\pi i}\int e^{\alpha(E-E')}\,d\alpha \quad ; \quad \delta_{NN'} = \frac{1}{2\pi i}\oint \frac{dz}{z^{N-N'+1}} = \frac{1}{2\pi i}\int e^{-\nu(N-N')}\,d\nu$$

Then:

$$W(E,N) = \frac{\epsilon}{(2\pi i)^2}\iint e^{\alpha E - \nu N + \log \Xi}\,d\nu\,d\alpha$$

We find the saddle point as before, it being given by:

$$E = -\frac{\partial}{\partial\alpha}\log\Xi \quad ; \quad N = \frac{\partial}{\partial\nu}\log\Xi \quad ; \quad \beta = \alpha\,, \quad \nu = \beta\mu$$

and:

$$\rho^{(1)}_{N_1E_1} = \frac{e^{-\beta(E_1-\mu N_1)}}{\Xi_1(\beta,\mu)}$$

Hence the density operator for subsystem 1 is:

$$\rho^{(1)} = \sum_{N_1E_1k} |N_1E_1k\rangle \frac{e^{-\beta(E_1-\mu N_1)}}{\Xi_1} \langle N_1E_1k|$$

We now define an operator N_1 such that $N_1|E N_1 k\rangle = N_1|E N_1 k\rangle$. On the left N_1 is an operator, on the right an eigenvalue. Thus we can write:

$$\boxed{\begin{gathered}\rho^{(1)} = \frac{e^{-\beta(H_1-\mu N_1)}}{\Xi_1(\beta,\mu)} \\[2em] \Xi_1(\beta,\mu) = \text{Trace } e^{-\beta(H_1-\mu N_1)}\end{gathered}}$$

We now apply what has been developed to the case of a non-interacting free particle gas. However, what results will have general application.

Quantum statistics of Free Particle Gases

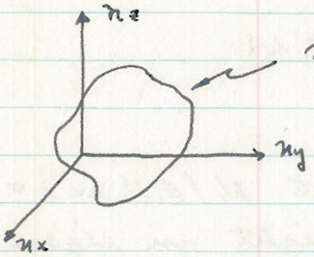
We place our systems in a cube of side L, volume L^3. We take periodic boundary conditions so that we have the convenience of working with running waves instead of traveling waves. The resulting plane waves are of the form:

$$e^{i\vec{k}\cdot\vec{r}} \; ; \quad k_x = \frac{2\pi}{L} n_x \,, \text{ etc} \; ; \quad n_x = 0, \pm 1, \pm 2, \cdots$$

We note that the k space is discreet: $\vec{k} = \frac{2\pi}{L}(n_x, n_y, n_z)$. We also have the usual energy relation:

$$E = \frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2) = \frac{\hbar}{2m}\left(\frac{2\pi}{L}\right)^2 |\vec{n}|^2$$

Consider the density of states by considering a volume in $\vec{n}$ space:



number of states is the volume of this space.

$$\therefore \; \Delta n = \left(\frac{L}{2\pi}\right)^3 \Delta\vec{k} = \frac{V}{(2\pi)^3}\Delta\vec{k}$$

$$= \frac{V}{(2\pi\hbar)^3}\Delta\vec{p} = \frac{V}{h^3}\Delta\vec{p}$$

using $\vec{p} = \hbar\vec{k}$. Thus, in a given volume of momentum space, there are $V/h^3 \, \Delta\vec{p}$ states.

We now make some qualitative remarks about the two types of quantum gases.

For both of these gases, we have markedly non-classical behaviour at low temperatures.

For the Fermi gas, we form a sphere in $\vec{p}$ space as this is the lowest energy configuration, that is, the fermions fill up from the lowest available state, but at the rate of only one fermion per state as a result of the PEP. The volume of the sphere can be written in terms of the total number of fermions:

$$(2S+1)\,\frac{4\pi}{3}(n)^3 = N \; ; \quad |n| = \left\{\frac{3N}{4\pi(2S+1)}\right\}^{1/3} \; ; \quad S = \text{particle spin.}$$

We then have: $k_F = \frac{2\pi}{L}|n| = \frac{2\pi}{L}\left\{\frac{3N}{4\pi(2s+1)}\right\}^{1/3}$; $p_F = \hbar k_F$

Finally:

$$p_F^3 = \frac{6\pi^2 N \hbar^3}{V(2s+1)} = \frac{6\pi^2\hbar^3}{2s+1}\left(\frac{1}{v}\right)$$

The fermion gas only exhibits classical behaviour when the particles obtain a temperature of $kT \gg E_F = \frac{p_F^2}{2m}$ so that the gas becomes rarefied and competition for states on behalf of the PEP becomes negligible.

For the Bose gas, all the particles go into the lowest state, but we still get non-classical behaviour except at high temperatures.

We now attempt to find a relation between the quantum and classical cases, that is, a correspondence relation. Recall that $W(E)$ is much like $\Omega(E)$. $W(E)$ is the number of states in energy ϵ, while $\Omega(E)$ is the volume of p.s per unit energy. Therefore the volume of p.s. in the interval ϵ is $\epsilon\Omega$. We can then write the correspondence relation:

$$W(E) = \frac{\epsilon\,\Omega(E)}{N!\,h^{3N}}$$

The $N!$ counteracts configurations that do not give different states.

For the partition functions, we have:

$$Q = \sum W e^{-\beta E} = \frac{1}{N!\,h^{3N}}\sum \epsilon\, e^{-\beta E}\,\Omega(E)$$

$$= \frac{\chi_{\text{classical}}(\beta, N)}{N!\ h^{3N}}\left\{1 + O(\hbar)\right\}$$

$$\Xi(\beta, z) = \sum z^N Q(\beta, N) = \sum \frac{z^N \chi_{\text{classical}}(\beta, N)}{N!\ h^{3N}}$$

$$= \Xi\left(\beta, \frac{z}{h^3}\right)\left\{1 + O(\hbar)\right\}$$

LECTURE 28 : 1-18-62

We are considering the quantum statistics of the free particle gas. The bookkeeping on the states of the gas can be done either by product (symmetric or antisymmetric) wave functions or by the use of the occupation number representation. The latter is more convenient and we will use that. Suppose l labels the single particle states with energy ϵ_l. The N particle system is then described by a set $\{n_l\}$ where n_l is the number of particles in state l. However we must admit restrictions on n_l for the two types of gases of the form:

$$\text{FD}: \quad n_l = 0, 1 \quad ; \quad \sum_l n_l = N$$

$$\text{EB}: \quad n_l = 0, 1, 2, \cdots \quad \text{and the total may be unrestricted.}$$

We denote a state of the occupation number representation by $|\{n_j\}\rangle$ so that it satisfies a Schroedinger equation of the form:

$$H|\{n_j\}\rangle = \sum_l n_l \epsilon_l \, |\{n_j\}\rangle$$

and we can form the partition function accordingly by taking the trace in the occupation number representation subject to the required constraint:

$$\boxed{Q(\alpha, N) = \sum_{\{n_l\}} \delta_{N, \Sigma n_l} \, e^{-\alpha \Sigma n_l \epsilon_l}}$$

We will first obtain the partition function (based on the CDF as above) without the constraint. This is the case for photons as their numbers are not conserved:

$$Q(\alpha) = \sum_{\{n_l\}} e^{-\alpha \Sigma n_l \epsilon_l} = \prod_l \sum_{n_l = 0}^{\infty} e^{-\alpha n_l \epsilon_l}$$

$$\boxed{Q(\alpha) = \prod_l \frac{1}{1 - e^{-\alpha \epsilon_l}}}$$

$\alpha = \beta$ of course, via the saddle point method.

For the internal energy, we have:

$$U = -\frac{\partial}{\partial\beta}\log Q = -\frac{\partial}{\partial\beta}\left\{-\sum_l \log\left(1-e^{-\beta\epsilon_l}\right)\right\}$$

$$= \sum_l \frac{\epsilon_l e^{-\beta\epsilon_l}}{1-e^{-\beta\epsilon_l}} = \sum_l \frac{\epsilon_l}{e^{\beta\epsilon_l}-1}$$

We can also find the average occupation of a state $\bar{n}_l$ from $Q_l(\beta)$:

$$\bar{n}_l = -\frac{1}{\beta}\frac{\partial}{\partial\epsilon_l}\log Q_l = \frac{e^{-\beta\epsilon_l}}{1-e^{-\beta\epsilon_l}} = \frac{1}{e^{\beta\epsilon_l}-1}$$

which is the familiar BE law for photons. Also we then see:

$$U = \sum_l \bar{n}_l\,\epsilon_l$$

For the variation of n_l, we get:

$$\overline{(n_l-\bar{n}_l)^2} = \frac{1}{\beta^2}\frac{\partial^2}{\partial\epsilon_l^2}\log Q_l = -\frac{1}{\beta}\frac{\partial}{\partial\epsilon_l}\bar{n}_l$$

$$= \frac{e^{\beta\epsilon_l}}{\left(e^{\beta\epsilon_l}-1\right)^2} = \frac{1}{e^{\beta\epsilon_l}-1} + \frac{1}{\left(e^{\beta\epsilon_l}-1\right)^2}$$

$$= \bar{n}_l^2 + \bar{n}_l = \bar{n}_l\left(\bar{n}_l+1\right)$$

This result is characteristic of the BE distribution, classically, it would be $\bar{n}_l$.

When the number of particles is conserved, it is easiest to use the GCDF. We could use the constraint $\delta_{N,\Sigma n_l}$ but to do so would involve much calculation. Also, it is more realistic to expect that there will be some haziness in the number of particles, thus making the GCDF more fitting to use.

$$\Xi(\alpha, z) = \sum_N z^N Q(\alpha, N)$$

Recall that we can invert this by Cauchy's Integral Formula: $Q(\alpha, N) = \frac{1}{2\pi i}\oint \frac{\Xi(\alpha, z)}{z^{N+1}}\,dz$

Thus if we can find the GPF $\overline{\underline{X}}$, we can also get Q. Proceeding, we plug in our previous result for Q and get:

$$\overline{\underline{X}} = \sum_N z^N \sum_{\{n_l\}} \delta_{N,\Sigma n_l} \prod_l e^{-\alpha n_l \epsilon_l}$$

$$= \sum_{\{n_l\}} \prod_l z^{n_l} e^{-\alpha n_l \epsilon_l} = \prod_l \sum_{n_l=0}^{UL} z^{n_l} e^{-\alpha n_l \epsilon_l}$$

where:

$$UL = \begin{cases} 1 : FD \\ \infty : EB \end{cases}$$

Hence we can immediately write:

$$FD: \quad \overline{\underline{X}} = \prod_l (1 + z e^{-\alpha \epsilon_l})$$

$$EB: \quad \overline{\underline{X}} = \prod_l \frac{1}{1 - z e^{-\alpha \epsilon_l}}$$

We may write these both together, letting $z = e^{\nu}$:

$$\boxed{\overline{\underline{X}}(\alpha, \nu) = \prod_l \left(1 \pm e^{\nu - \alpha \epsilon_l}\right)^{\pm 1} \qquad \begin{matrix} + \rightarrow FD \\ - \rightarrow EB \end{matrix}}$$

The sddle point is given by:

$$n = \bar{N} = \frac{\partial}{\partial \nu} \log \overline{\underline{X}} \quad ; \quad U = \bar{E} = -\frac{\partial}{\partial \alpha} \log \overline{\underline{X}}$$

given $\alpha = \beta$; $\nu = \beta \mu$. Now:

$$\frac{\partial}{\partial \nu} \log \overline{\underline{X}} = \frac{\partial}{\partial \nu} \sum_l \pm \log (1 \pm e^{\nu - \alpha \epsilon_l})$$

$$= \sum_l \frac{e^{\nu - \alpha \epsilon_l}}{1 \pm e^{\nu - \alpha \epsilon_l}}$$

Hence:

$$\boxed{n = \bar{N} = \sum_l \frac{1}{e^{\beta(\epsilon_l - \mu)} \pm 1}}$$

And:

$$\boxed{U = \sum_l \frac{\epsilon_l}{e^{\beta(\epsilon_l - \mu)} \pm 1}}$$

Also.
$$\overline{n_l} = \frac{1}{\beta}\frac{\partial}{\partial \epsilon_l} \log \mathfrak{X}_l = \frac{1}{e^{\beta(\epsilon_l-\mu)} \pm 1}$$

and:
$$N = \sum_l \overline{n_l} \quad ; \quad U = \sum_l \overline{n_l}\, \epsilon_l$$

The above are more or less general for non-interacting quantum gas systems. We now consider some special cases.

The Photon Gas

We take the radiation to be composed of plane waves subject to the BVK boundary conditions:

$$e^{i\vec{k}\cdot\vec{r}} \quad ; \quad \vec{k} = \frac{2\pi}{L}(n_x, n_y, n_z) \quad ; \quad d\vec{n} = \frac{V}{(2\pi)^3}\, d\vec{k}$$

now, the number of modes in $d\vec{n} = \frac{2V}{(2\pi)^3}\, d\vec{k}$, the extra 2 coming from polarization.

We have:

$$\log Q = -\sum_l \log\left(1 - e^{-\beta\epsilon_l}\right)$$

$$\rightarrow \frac{-2V}{(2\pi)^3}\int \log\left(1 - e^{-\beta\hbar c k}\right) d\vec{k}$$

$$= -\frac{8\pi}{(2\pi)^3} V \int_0^\infty \log\left(1 - e^{-\beta\hbar c k}\right) k^2\, dk$$

$$= -\frac{8\pi}{(2\pi)^3} V \left\{ \frac{k^3}{3}\log\left(1 - e^{-\beta\hbar c k}\right)\Big|_0^\infty - \frac{\beta\hbar c}{3}\int_0^\infty \frac{k^3\, dk}{e^{\beta\hbar c k} - 1}\right.$$

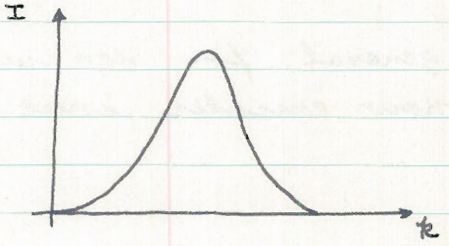
$$= \frac{8\pi V}{3(2\pi)^3}\beta\hbar c \int_0^\infty \frac{k^3\, dk}{e^{\beta\hbar c k} - 1}$$

$$F = -kT\log Q = -\frac{8\pi V}{3(2\pi)^3}\hbar c\int_0^\infty \frac{k^3\, dk}{e^{\beta\hbar c k} - 1}$$

Also: $U = -\frac{\partial}{\partial\beta}\log Q = \sum_{k} \frac{\epsilon_k}{e^{\beta\epsilon_k}-1} \rightarrow \frac{2V(4\pi)}{(2\pi)^3}\hbar c \int_0^\infty \frac{k^3\,dk}{e^{\beta\hbar c k}-1}$

or: $\boxed{F = -\frac{1}{3}U}$

The plot of the integrand versus k gives Planck's Law of Black Body Spectral Density:



Since the electromagnetic field is a collection of harmonic oscillators, we would expect by the classical equipartition theorem, that the internal energy of the field is given by:

$$U = \frac{2V}{(2\pi)^3}\int kT\,d\vec{k} = \frac{8\pi V}{(2\pi)^3}kT\int_0^\infty k^2\,dk$$

The integrand expresses the Rayleigh-Jeans Law which holds for low frequencies but blows up for higher frequencies. Planck first derived his law by interpolation between the observed high and low frequency behaviour.

What is the equation of state of the field?

$$p = -\frac{\partial F}{\partial V} = \frac{1}{3}\frac{\partial U}{\partial V}$$

or: pressure = $\frac{1}{3}$ energy density

We now attempt to evaluate the integral that appears in F and U. Let:

$$x = \beta\hbar c k = \frac{\hbar\omega}{kT}$$

Then we have to consider:

$$\int_0^\infty \frac{k^3\,dk}{e^{\beta\hbar c k}-1} = \frac{1}{(\beta\hbar c)^4}\int_0^\infty \frac{x^3\,dx}{e^x-1}$$

Consider:

$$\int_0^\infty \frac{x^{n-1}}{e^x-1}\,dx = \int_0^\infty x^{n-1} e^{x} \frac{dx}{1-e^{-x}} = \int_0^\infty x^{n-1} \sum_{j=1}^{\infty} e^{-jx}\,dx$$

$$= \Gamma(n) \sum_{j=1}^{\infty} \frac{1}{j^n} = \Gamma(n)\,\zeta(n)$$

where $\zeta(n)$ = Riemann's Zeta Function. For n even, we have the following relation to the Bernoulli numbers:

$$\zeta(n) = \frac{(2\pi)^n}{(2n)!} B_{n/2} \quad ; \qquad \begin{array}{ll} B_1 = \frac{1}{6} \; ; & B_2 = \frac{1}{30} \\ B_3 = \frac{1}{42} \; ; & B_4 = \frac{1}{30} \end{array}$$

Hence for n even, we have:

$$\int_0^\infty \frac{x^{n-1}}{e^x-1}\,dx = \frac{(2\pi)^n}{2n} B_{n/2}$$

and:

$$\int_0^\infty \frac{x^3\,dx}{e^x-1} = \frac{\pi^4}{15}$$

Therefore we have:

$$\boxed{\begin{array}{l} \log Q = \dfrac{\pi^2}{45}\left(\dfrac{kT}{\hbar c}\right)^3 V \quad ; \quad F = -\dfrac{\pi^2}{45}\dfrac{(kT)^4}{(\hbar c)^3} V \\[2ex] U = \dfrac{\pi^2}{15}\dfrac{(kT)^4}{(\hbar c)^3} V \quad (\text{Stephan's Law}) \end{array}}$$

The Fermi Gas

The number of states per unit energy is given by, using $p = \hbar k$:

$$g \frac{V}{(2\pi)^3 \hbar^3} 4\pi p^2 \frac{dp}{dE}$$

where $g = 2S+1$, S = spin. The number of states per unit energy can also be written in terms of the classical structure function for one particle:

$$g\,\frac{\Omega(E)}{h^3} = \frac{g}{h^3}\, 2\pi\, (2m)^{3/2} \epsilon^{1/2}\, V$$

In the following development, we will also include the EB case for a while:

$$\log \Xi = \log \prod_{l} \left(1 \pm e^{\beta(\mu - \epsilon_l)}\right)^{\pm 1}$$

$$= \pm \sum_{l} \log\left(1 \pm e^{\beta(\mu - \epsilon_l)}\right)$$

Then:

$$\log \Xi = \pm \frac{g}{h^3} (2\pi)(2m)^{3/2} V \int_0^\infty \epsilon^{1/2} \log\left(1 \pm e^{\beta(\mu-\epsilon)}\right) d\epsilon$$

$$= \frac{2}{3} \beta \frac{g}{h^3} (2\pi)(2m)^{3/2} V \int_0^\infty \frac{\epsilon^{3/2}\, d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1}$$

Also:

$$\boxed{N = \sum_{l} \frac{1}{e^{\beta(\epsilon_l - \mu)} \pm 1} = \frac{g}{h^3}(2\pi)(2m)^{3/2} V \int_0^\infty \frac{\epsilon^{1/2}\, d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1}}$$

This is a very general relation for μ in terms of N but it is hard to invert. Likewise we have:

$$\boxed{U = \sum_{l} \frac{\epsilon_l}{e^{\beta(\epsilon_l - \mu)} \pm 1} = \frac{g}{h^3}(2\pi)(2m)^{3/2} V \int_0^\infty \frac{\epsilon^{3/2}\, d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1}}$$

Note that: $\log \Xi = \frac{2}{3} \beta U$

Recall: $J = F - \mu N = U - TS - \mu N = -kT \log \Xi = -\frac{2}{3} U$

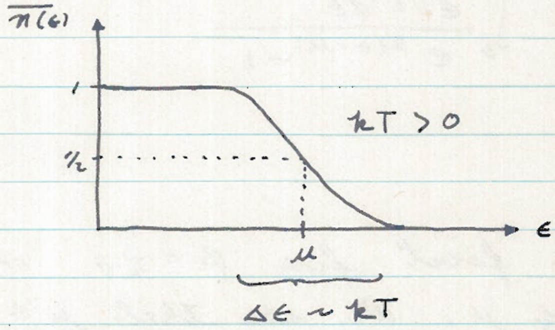
and for the case of a homogeneous system, $J = -pV$. Thus we have the result for both FD and EB particles:

$$\boxed{pV = \frac{2}{3} U \quad ; \quad p = \frac{1}{\beta}\frac{\partial}{\partial V} \log Q = \frac{2}{3}\frac{U}{V}}$$

now classically and for high temperatures $U = \frac{3}{2} NkT$ but in general $U \neq \frac{3}{2} NkT$, so above is most general equation of state for gases.

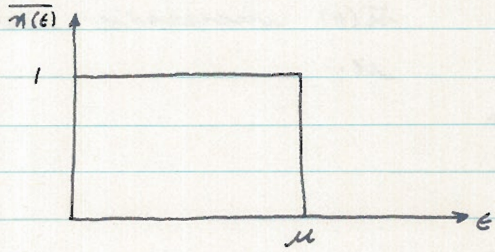
Returning to the case of a Fermi gas, it is hard to find U analytically unless we consider the gas at low temperature. We plot what is commonly known as the Fermi distribution law:

$$\overline{n}(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)}+1}$$



We see that the only active part of a Fermi gas is in a width kT about the Fermi surface.

Now for $T = 0$:



$$\overline{n}(\epsilon) = \begin{cases} 1 & \epsilon < \mu \\ 0 & \epsilon > \mu \end{cases}$$

Under these conditions, we have for N an analytic expression:

$$N = \frac{g}{h^3} 2\pi (2m)^{3/2} V \int_0^{\mu} \epsilon^{1/2}\, d\epsilon = \frac{g}{h^3} \frac{4\pi}{3} (2m\mu)^{3/2} V$$

Then: $g \frac{4\pi}{3} \frac{p_F^3 V}{h^3} = N$; $\mu = \epsilon_F$ at $T = 0$

$$U = \frac{g}{h^3} (2\pi)(2m)^{3/2} V \int_0^{\epsilon_F} \epsilon^{3/2}\, d\epsilon = \frac{g}{h^3} 2\pi (2m)^{3/2} V \frac{2}{5} \epsilon_F^{5/2}$$

$$= \frac{g}{h^3} \frac{4\pi}{5} \frac{V}{2m} p_F^5$$

or:

$$U = \frac{g}{h^3} \left(\frac{3h^3}{4\pi g}\right)^{5/3} \frac{4\pi}{5} \frac{1}{2m} N \left(\frac{N}{V}\right)^{2/3}$$

The equation of state of the gas at $T=0$ is:

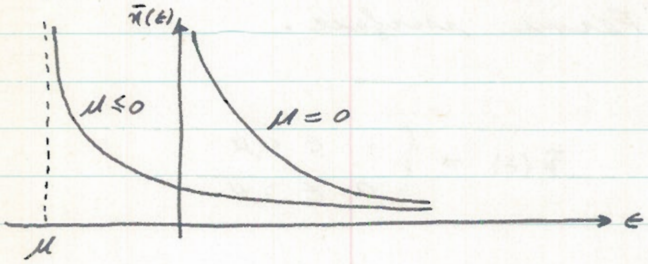
$$\boxed{p = \frac{2}{3}\frac{U}{V} = \frac{g}{h^3}\left(\frac{3h^3}{4\pi g}\right)^{5/3}\frac{8\pi}{15}\frac{1}{2m}\left(\frac{N}{V}\right)^{5/3}}$$

The Bose-Einstein Gas

$$N = \frac{g}{h^3}(2\pi)(2m)^{3/2}V\int_0^\infty \frac{\epsilon^{1/2}\,d\epsilon}{e^{\beta(\epsilon-\mu)}-1}$$

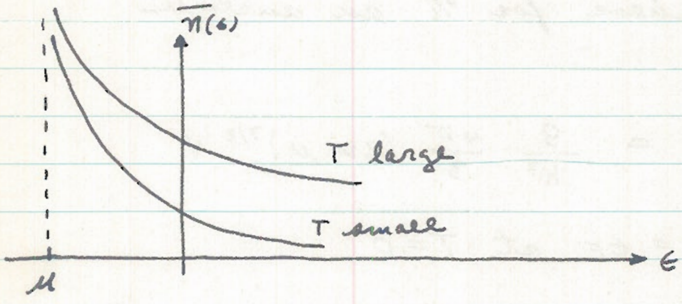
$$\overline{n(\epsilon)} = \frac{1}{e^{\beta(\epsilon-\mu)}-1}$$

We now plot $\overline{n(\epsilon)}$. We first fix $\beta = \frac{1}{kT}$ and plot for various values of μ. We see that $\mu \leq 0$ because we would otherwise have singlarties occuring for positive values of energy:



We see that for fixed T, $\overline{n}(\epsilon)$ increases with increasing μ.

Now fix μ and vary T:



Then for fixed μ, $\overline{n}(\epsilon)$ increases with T.

If we take $\overline{n(\epsilon)}$ fixed and let T decrease, we see that μ must increase to keep $\overline{n(\epsilon)}$ fixed. That is, $\mu \to 0$.

We may now write out N for $\mu = 0$.

$$N = \frac{g}{h^3} (2\pi)(2m)^{3/2} V (kT)^{3/2} \int_0^\infty \frac{x^{1/2}\,dx}{e^x - 1}$$

$$= \frac{g}{h^3}\, 2\pi (2mkT)^{3/2} V\, \Gamma\left(\tfrac{3}{2}\right) J\left(\tfrac{3}{2}\right) \quad ; \quad J\left(\tfrac{3}{2}\right) = 2.612$$

$$\frac{N}{V} = \frac{g}{h^3} \pi^{3/2} J\left(\tfrac{3}{2}\right) (2mkT)^{3/2}$$

Now this relation defines a minimum or critical temperature T_c. We now know that for $T > T_c$, $\mu < 0$, but what about $T < T_c$? If we fix $T > T_c$, increase N until $\mu = 0$ it would seem that at any given temperature we would have an upper limit on N which is nonsense. The trouble lies in the use of $\int$ instead of Σ. If we had used Σ we would have escaped this paradox. The thing is that at low enough temperatures nearly all of the particles can be in the ground state while for fermions this is not possible. Consider $\overline{n}_0$ the average number in the ground state:

$$\overline{n}_0 = \frac{1}{e^{-\beta\mu} - 1} \gg 1$$

and for $T < T_c$ we can have $\overline{n}(0) = O(N)$, that is, almost all the bosons will be in the ground state. at the critical temperature T_c we then say that we have Bose-Einstein Condensation. Let us separate N into two parts:

$$N = N_0 + N'$$

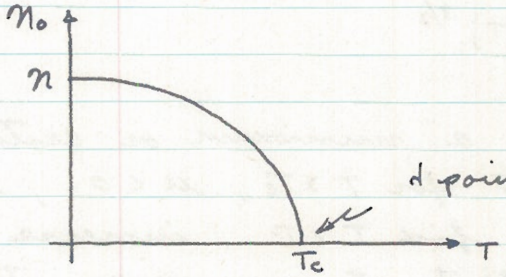
and say that it is N' that is given above for $\mu = 0$, that is:

$$N' = \frac{g}{h^3} \pi^{3/2} (2mkT)^{3/2} V J\left(\tfrac{3}{2}\right)$$

and: $$\frac{N'}{N} = \left(\frac{T}{T_c}\right)^{3/2}$$

Then: $n_0 = n - n' = n\left(1 - \frac{n'}{n}\right) = n\left[1 - \left(\frac{T}{T_c}\right)^{3/2}\right]$

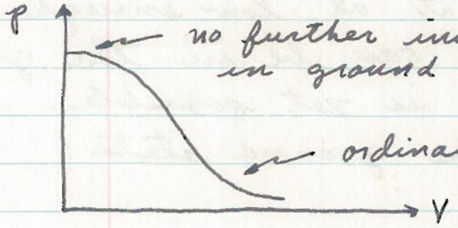
If we put in the appropriate numbers for Helium, we can obtain the following curve, knowing the λ point for liquid He or the superfluid point.



λ point, λ being the de Broglie wavelength of kT_c.

We see that below T_c, the ground state rapidly fills up.

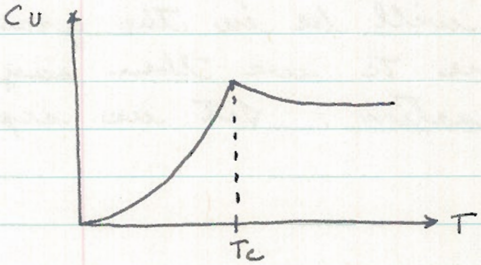
The isotherms are:



no further increase in pressure because all are in ground state.

ordinary isotherm

The specific heat shows a marked anomaly at T_c



This concludes the formal lectures.

Physics 262
Final Examination
January 27, 1962

Answer all three questions. Credits for the various sections are shown in parentheses.

1. (a) (5) Define the structure function of a classical mechanical system.

   (b) (10) A particle of mass m moves inside a rectangular box in a uniform vertical gravitational field. The box has length ℓ, width w and height h. Find the structure function for the system.

   (c) (10) Find the partition function for the particle in the box in a gravitational field.

   (d) (10) Assume that the box described above contains a gas of N identical non-interacting particles at temperature T. Find the variance of the height of the center of gravity of the gas particles for the limit in which the height of the box becomes infinite.

2. (a) (15) State and prove the classical equipartition theorem, and evaluate the expressions involved, wherever possible, for the microcanonical distribution and the canonical distribution.

   (b) (5) Find the heat capacity of a collection of N identical independent three-dimensional (isotropic) harmonic oscillators of mass m and and vibration frequency $\omega/2\pi$.

   (c) (15) In one of the earliest models of the thermal agitation of atoms in solids the atoms were regarded as quantum mechanical harmonic oscillators analogous to those described in (b). Find expressions for the heat capacity of the quantum mechanical system valid in the limits of low temperature and of high temperature.

3. A gas of N identical non-interacting particles of mass m and spin 1/2 occupies a volume V.

   (a) (15) Use the grand canonical ensemble to derive the energy distribution of the particles as a function of temperature, T.

   (b) (5) State relations valid at all temperatures from which the internal energy U and the pressure P may be evaluated.

   (c) (10) Evaluate U and P at temperature T = 0 in terms of the above constants.

Paul M. Grant
Physics 262

Assignment No. 1

Note to grader: I thought this assignment was due Thursday; that is why it is late.

1
10
10
10
10
6

47/60

P262 : Problem 1

Consider $G(q_1 \cdots p_F, t)$ and $\bar{G}(t)$ and $\frac{d}{dt}\bar{G}(t)$.

show that:

$$\frac{d}{dt}\bar{G}(t) = \overline{\frac{dG}{dt}}$$

where $\frac{dG}{dt}$ is a derivative which follows motion thru a phase point. Consider $\rho(q_1 \cdots p_F, t)$ depending on everything.

Recall the definition:

$$\bar{G}(t) = \int G(q_1 \cdots p_F, t)\, \rho(q_1 \cdots p_F, t) \prod_{j=1}^{F} dq_j\, dp_j$$

Form by direct differentiation:

$$\frac{d\bar{G}(t)}{dt} = \int \underset{\left(\frac{\partial G}{\partial t}\right)}{\frac{dG(q_1 \cdots p_F, t)}{dt}}\, \rho(q_1 \cdots p_F, t) \prod_{j=1}^{F} dq_j\, dp_j$$

$$+ \int G(q_1 \cdots p_F, t)\, \underset{\left(\frac{\partial \rho}{\partial t}\right) = -v \cdot \nabla \rho}{\frac{d\rho(q_1 \cdots p_F, t)}{dt}} \prod_{j=1}^{F} dq_j\, dp_j$$

However, by the Liouville theorem, which states:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + [\rho, H] = 0$$ Yes but . . . .

we have immediately:

$$\frac{d\,\bar{g}(t)}{dt} = \int \frac{d\,g(q_1 \cdots p_F, t)}{dt}\, \rho(q_1 \cdots p_F, t) \prod_{j=1}^{F} dq_j\, dp_j$$

or:

$$\boxed{\frac{d\,\bar{g}(t)}{dt} = \overline{\left(\frac{d\,g(t)}{dt}\right)}} \qquad ①$$

$$\cancel{\frac{d}{dt} f(x,t)}$$

$$\cancel{f(t) = \int g(x,t)\,dx}$$

$$\cancel{\frac{df(t)}{dt} = \int \left(\frac{\partial}{\partial x} g(x,t)\right)}$$

$$f(x) = \int dy\, g(x,y)$$

$$\frac{d}{dx} f = \int dy \left(\frac{\partial}{\partial x} g(x\,y)\right)_y$$

Integrations by parts are needed.

P262 : Problem 2

Consider the generation of random binary numbers, N digits long. The a priori probability of any given one is $\frac{1}{2^N}$ since there are 2^N possible numbers.

Label the digits of a number by a_j, $j = 1, \ldots, N$; a_j being either 0 or 1. Define the average of the digits in any given number by:

$$A = \frac{1}{N} \sum_{j=1}^{N} a_j$$

which will of course be very close to 1/2 if N is large. Consider the deviation from 1/2 and find the probability distribution around 1/2. A way of doing this is to consider $N \gg 1$ and change the Σ to $\int$. Thus show the probability density about A is:

$$P(A) = \sqrt{\frac{2N}{\pi}}\, e^{-2N(A-1/2)^2}$$

in the asymptotic limit of N large and continuous. This could be done with the binomial coefficient and Stirling's approximation and integrating. Also could be done with the central limit theorem.

Now consider $N = 10^{20}$: find to one significant figure the fraction of points in phase space for which $|A - 1/2| > 10^{-8}, 10^{-9}, 10^{-10}$.

We assume that there will be $\frac{1}{2}(N+n)$ of the digits in the state 1 and of course $\frac{1}{2}(N-n)$ in the state zero. The average value of a digit is then obviously:

$$A = \frac{1}{N}\sum_{j=1}^{N} a_j = \frac{N+n}{2N} = \frac{1}{2} + \frac{n}{2N}$$

$$\text{or} \quad \frac{n}{2N} = A - \frac{1}{2}$$

since a_j exists only for values of 1. It follows from combinatorial procedures that the arrangement of the groups of $\frac{1}{2}(N+n)$ and $\frac{1}{2}(N-n)$ numbers among the N digits is:

$$\frac{N!}{\left\{\frac{1}{2}(N+n)\right\}!\left\{\frac{1}{2}(N-n)\right\}!}$$

thus giving the number of indistinguishable arrangements. Now the a-priori probability of a giving arrangement is $\left(\frac{1}{2}\right)^N$. Thus we have:

$$P_N(n) = \left(\frac{1}{2}\right)^N \frac{N!}{\left\{\frac{1}{2}(N+n)\right\}!\left\{\frac{1}{2}(N-n)\right\}!}$$

$$\log P_N(n) = -N\log 2 + \log N!$$

$$-\log\left\{\frac{N}{2}\left(1+\frac{n}{N}\right)\right\}! - \log\left\{\frac{N}{2}\left(1-\frac{n}{N}\right)\right\}!$$

We use Stirlings approximation: $N \gg 1; \gg n$

$$\log(n-1)! = (n - 1/2)\log n - n + \frac{1}{2}\log 2\pi$$

: $$\log n! = (n+\tfrac{1}{2})\log(n+1) - n - 1 + \tfrac{1}{2}\log 2\pi$$

$$\log P_N(n) = -N\log 2 + (N+\tfrac{1}{2})\log N(1+\tfrac{1}{N}) \underline{-N} \underline{-1 + \tfrac{1}{2}\log 2\pi}$$

$$-\left(\frac{N}{2}(1+\frac{n}{N})+\frac{1}{2}\right)\log\left(\frac{N}{2}(1+\frac{n}{N})+1\right) + \underline{\frac{N}{2}(1+\frac{n}{N})} \underline{+1} \underline{-\frac{1}{2}\log 2\pi}$$

$$-\left(\frac{N}{2}(1-\frac{n}{N})+\frac{1}{2}\right)\log\underbrace{\left(\frac{N}{2}(1-\frac{n}{N})+1\right)}_{\frac{N}{2}-\frac{n}{2}+1} + \underline{\frac{N}{2}(1-\frac{n}{N})} + 1 - \frac{1}{2}\log 2\pi$$

$$= \frac{N}{2} - \frac{n+2}{2} = \frac{N}{2}\left(1 - \frac{n+2}{N}\right)$$

Use $$\log\left(1 \pm \frac{n+2}{N}\right) \simeq \pm\frac{n+2}{N} - \frac{(n+2)^2}{2N^2} + \cdots$$

Then: $$\log P_N(n) = -N\log 2 + (N+\tfrac{1}{2})\left[\log N + \frac{1}{N} - \frac{1}{2N^2}\right]$$

$$-\left(\frac{N}{2}+\frac{n}{2}+\frac{1}{2}\right)\left[\log N - \log 2 + \frac{n+2}{N} - \frac{(n+2)^2}{2N^2}\right]$$

$$-\left(\frac{N}{2}-\frac{n}{2}+\frac{1}{2}\right)\left[\log N - \log 2 - \frac{n+2}{N} - \frac{(n+2)^2}{2N^2}\right] + 1 - \frac{1}{2}\log 2\pi$$

Collecting terms:

$$\log P_N(n) = +\frac{1}{2}\log N + \frac{1}{2}\log 4 - \frac{1}{2}\log 2\pi - \frac{(n+2)^2}{2N}$$

But $$\frac{(n+2)^2}{2N} \cong \frac{n^2}{2N} \quad \text{for} \quad n \ll N$$

Now: $\frac{n^2}{2N} = 2N\left(\frac{n}{2N}\right)^2 = 2N\left(A - 1/2\right)^2$

Then:

$$\boxed{P_N(A) = \sqrt{\frac{2N}{\pi}}\; e^{-2N(A-1/2)^2}}$$

The maximum value of A is 1 and the minimum 0.

$|A - 1/2|_{max} = 1/2$ or: $-1/2 < A - 1/2 < 1/2$,

Now:

$$P_N\left(|A-1/2| > x\right) = 2\sqrt{\frac{2N}{\pi}} \int_x^{1/2} e^{-2Ny^2}\, dy$$

Let $z = \sqrt{2Ny^2}$, $dy = \frac{dz}{\sqrt{2N}}$

Then: $$P_N\left(|A-1/2| > x\right) = \frac{2}{\sqrt{\pi}} \int_{\sqrt{2N}\,x}^{\sqrt{\frac{N}{2}}} e^{-z^2}\, dz$$

$$= \frac{2}{\sqrt{\pi}} \left\{ \int_0^{\sqrt{\frac{N}{2}}} e^{-z^2} dz - \int_0^{\sqrt{2N}\,x} e^{-z^2} dz \right\}$$

Now: $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-r^2} dr$, Then:

$$P_N\left(|A-1/2| > x\right) = \operatorname{erf}\left(\sqrt{\frac{N}{2}}\right) - \operatorname{erf}\left(\sqrt{2N}\,x\right)$$

For large x, the asymptotic expression is:

$$\operatorname{erf}(x) \sim 1 - \frac{1}{\sqrt{\pi}\,x} e^{-x^2}$$

Since $N = 10^{20}$, $\operatorname{erf}(10^{10}) = 1$

so:

$$P_{10^{20}}\left(|A - 1/2| > x\right) \approx \frac{e^{-x^2 \cdot 2N}}{\sqrt{\pi}\,\sqrt{2N}\,x}$$

$$P_{10^{20}}\left(|A - 1/2| > x\right) = \frac{e^{-20000}}{100\sqrt{2\pi}} \quad ; \quad x = 10^{-8}$$

[illegible] $\approx \left(10^{-.434}\right)^{20000}$

o.k.

$$P_{10^{20}}\left(|A - 1/2| > x\right) = \frac{e^{-200}}{10\sqrt{2\pi}} \quad ; \quad x = 10^{-9}$$

[illegible] $\approx \left(10^{-.434}\right)^{200}$

more exactly:

$$P_{10^{20}}\left(|A - 1/2| > x\right) = 1 - \operatorname{erf}\left(\sqrt{2N}\,x\right)$$

$$P_{10^{20}}\left(|A - 1/2| > 10^{-10}\right) = 1 - \operatorname{erf}\left(\sqrt{2}\right) = 1 - \operatorname{erf}(1.41)$$

$$= 1 - .9539 = .0461 \approx .05$$

✓

The only time an appreciable amount of the phase space is occupied is for $|A - 1/2| > 10^{-10}$. The error or fraction occupied for $|A - 1/2| > 10^{-8}$ is negligible. Shows tremendous convergence of the mean value for $N = 10^{20}$.

10

P262 : Problem 3

We suppose a perfect gas of $F = 3N$ momentum components. Choose the distribution function ρ not larger than a certain energy:

$$\rho(q_1 \cdots p_1 \cdots) = \text{constant} \times \Psi_E(\mathcal{H})$$

$$\Psi_E(\mathcal{H}) = \begin{cases} 0, & \mathcal{H} > E \\ 1, & \mathcal{H} \leq E \end{cases}$$

Find the distribution: $P(p_1)$

$$P(p_1) = \frac{\int \rho \prod_{j=2}^{F} dp_j \prod_{j=1}^{F} dq_j}{\int \rho \prod_{j=1}^{F} dp_j \prod_{j=1}^{F} dq_j}$$

$$= \frac{\int \rho \prod_{j=2}^{F} dp_j}{\int \rho \prod_{j=1}^{F} dp_j}$$ since the Hamiltonian is independent of q_j.

$$\mathcal{H} = \sum_{j=1}^{F} \frac{p_j^2}{2m}$$

$$P(p_1) = \frac{\int \Psi_{E - \frac{p_1^2}{2m}}\left(\mathcal{H} - \frac{p_1^2}{2m}\right) \prod_{j=2}^{F} dp_j}{\int \Psi_E(\mathcal{H}) \prod_{j=1}^{F} dp_j}$$

$$\int \Psi_{E-\frac{p_1^2}{2m}}\left(\mathcal{H}-\frac{p_1^2}{2m}\right)\prod_{j=2}^{F} dp_j$$

$$= \int\int_0^{E-\frac{p_1^2}{2m}} \delta\left(\left[E'-\frac{p_1^2}{2m}\right]-\sum_{j=1}^{F}\frac{p_j^2}{2m}+\frac{p_1^2}{2m}\right)\prod_{j=2}^{F} dp_j \; d\left(E'-\frac{p_1^2}{2m}\right)$$

$$= \int\int_0^{E-\frac{p_1^2}{2m}} \delta\left(\left[E'-\frac{p_1^2}{2m}\right]-\sum_{j=1}^{F-1}\frac{p_j^2}{2m}\right)\prod_{j=1}^{F-1} dp_j \; d\left(E'-\frac{p_1^2}{2m}\right)$$

$$= (2m)^{+\frac{F-1}{2}}\frac{A_{F-1}}{2}\int_0^{E-\frac{p_1^2}{2m}}\left(E'-\frac{p_1^2}{2m}\right)^{\frac{F-3}{2}} d\left(E'-\frac{p_1^3}{2m}\right)$$

$$= (2m)^{+\frac{F-1}{2}}\frac{A_{F-1}}{2}\,\frac{2}{F-1}\left(E-\frac{p_1^2}{2m}\right)^{\frac{F-1}{2}}$$

using the δ function representation of $\Psi_E(\mathcal{H})$.

$$\int \Psi_E(\mathcal{H})\prod_{j=1}^{F} dp_j = (2m)^{+\frac{F}{2}}\frac{A_F}{2}\,\frac{2}{F}\,E^{F/2}$$

Then:

$$P(p_1) = \frac{1}{\sqrt{2m}}\,\frac{A_{F-1}}{A_F}\,\frac{F}{F-1}\,\frac{\left(E-\frac{p_1^2}{2m}\right)^{\frac{F-1}{2}}}{E^{\frac{F}{2}}}$$

$$= \frac{1}{\sqrt{2mE}}\,\frac{A_{F-1}}{A_F}\,\frac{1}{1-1/F}\left(1-\frac{p_1^2}{2mE}\right)^{\frac{F-1}{2}}$$

Now for small p_1, large E and F, using

$$(1-x)^n \approx e^{-nx}$$

we have:

$$P(p_1) = \frac{A_{F-1}}{A_F}\,\frac{1}{\sqrt{2mE}}\,e^{-\frac{p_1^2}{2mE}\frac{F}{2}}$$

now: $A_n = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$, so:

$$\frac{A_{F-1}}{A_F} = \frac{2\pi^{\frac{F-1}{2}}}{\Gamma(\frac{F-1}{2})} \cdot \frac{\Gamma(\frac{F}{2})}{2\pi^{F/2}} = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{F}{2})}{\Gamma(\frac{F-1}{2})}$$

Recall from lecture:

$$\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)} \simeq \sqrt{n} \quad \text{for } n \gg 1$$

Then finally; using $F = 3N$

$$\boxed{P(p_1) = \sqrt{\frac{1}{2\pi m}\frac{3N}{2E}}\; e^{-\frac{p_1^2}{2m}\left(\frac{3N}{2E}\right)}}$$

✓ ⑩

Thus we obtain the Maxwell-Boltzmann distribution. Thus, even though we took $\rho(q_1 \cdots p_1 \cdots)$ as constant inside the hypersphere bounded by the constant energy surface, we get that the energy of the system will be practically E or that most of the volume of the hypersphere is contained very close to the surface of constant energy.

P 262 : Problem 4 :

ⓐ for a 1-D harmonic oscillator, find $\Omega(E)$:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{k}{2}q^2 = \frac{1}{2m}\left(p^2 + m^2\omega^2 q^2\right) \quad ; \quad k = m\omega^2$$

ⓑ Find $\Omega(E)$ for 3-D harmonic oscillator:

$$\mathcal{H} = \frac{1}{2m}\left(|p|^2 + m^2\omega^2 r^2\right)$$

ⓒ Find $\Omega(E)$ for rigid rotator, A = moment of inertia

$$\mathcal{H} = \frac{1}{2A}\left\{p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right\}$$

ⓐ $\Omega(E) = \int \delta\left(E - \mathcal{H}(q_1 \cdots p_F)\right) dV$

Recall the property of the δ function:

$$\delta(f(x)) = \frac{1}{|f'(x_0)|}\,\delta(x - x_0) \quad \text{where } x_0 \text{ is a root of } f(x)$$

Here:

$$\Omega(E) = \int \delta\left(E - \frac{1}{2m}\left(p^2 + m^2\omega^2 q^2\right)\right) dp\, dq$$

let: $x = \frac{p}{\sqrt{2m}}$; $y = \sqrt{\frac{m}{2}}\,\omega q$

Then: $\Omega(E) = \frac{2}{\omega}\int \delta\left(E - x^2 - y^2\right) dx\, dy$

But this is in the form $\int \delta\left(r^2 - \sum_1^n x_j^2\right) \prod_1^n dx_j = \frac{1}{2} r^{n-2} A_n$

where:

$$A_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}$$

Hence: $\Omega(E) = \frac{2}{\omega} \cdot \frac{1}{2} E^{0} \frac{2\pi}{\Gamma(1)} = \frac{2\pi}{\omega} = \frac{1}{f}$ ✓

or $\boxed{\Omega(E) = \text{period} = \frac{2\pi}{\omega}}$

ⓑ

$$\Omega(E) = \int \delta\left(E - \frac{1}{2m}\left[(p_x^2 + p_y^2 + p_z^2) + m^2\omega^2(x^2+y^2+z^2)\right]\right) dp_x\, dp_y\, dp_z\, dx\, dy\, dz$$

$$= \left(\frac{2}{\omega}\right)^3 \int \delta\left(E - \sum_1^6 x_j\right) \prod_1^6 dx_j = \left(\frac{2}{\omega}\right)^3 \cdot \frac{1}{2} \sqrt{E}^{\,4} A_6$$

$$= \left(\frac{2}{\omega}\right)^3 \cdot \frac{1}{2} \sqrt{E}^{\,4} \cdot 4\pi^3 = \frac{16\pi^3 E^2}{4\,\omega^3}$$

or: $\boxed{\Omega(E) = \frac{16\pi^3 E^2}{4\,\omega^3}}$ $\boxed{\Omega(E) = \frac{4\pi^3 E^2}{\omega^3}}$ ✓

ⓒ

$$\Omega(E) = \int \delta\left\{E - \frac{1}{2A}\left(p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}\right)\right\} dp_\theta\, dp_\phi\, d\phi\, d\theta$$

$$= 2\pi \int \delta\left\{E - \frac{1}{2A}\left(p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}\right)\right\} dp_\theta\, dp_\phi\, d\theta$$

The roots with respect to p_θ are at

$$p_\theta = \sqrt{2AE - \frac{p_\phi^2}{\sin^2\theta}}$$

$$f(p_\theta) = E - \frac{1}{2A} p_\theta^2 - \frac{1}{2A} \frac{p_\phi^2}{\sin^2\theta}$$

$$\frac{f(p_\theta)}{\partial p_\theta} = -\frac{p_\theta}{A} = -\frac{1}{A}\left\{2AE - \frac{p_\varphi^2}{\sin^2\theta}\right\}^{1/2}$$

Thus:

$$\Omega(E) = 4\pi A \int \frac{dp_\varphi\, d\theta}{\left[2AE - \frac{p_\varphi^2}{\sin^2\theta}\right]^{1/2}}$$

$$= 4\pi A \int dp_\varphi \int d\theta \, \frac{\sin\theta}{\left[2AE\sin^2\theta - p_\varphi^2\right]^{1/2}}$$

$$= \frac{4\pi A}{\sqrt{2AE}} \int dp_\varphi \int_0^\pi \frac{\sin\theta\, d\theta}{\left[\sin^2\theta - \frac{p_\varphi^2}{\cancel{\sin\theta}\ 2AE}\right]^{1/2}}$$

$$= 4\pi A \int_0^\pi d\theta\, \sin\theta \int_{-\sqrt{2AE}\,\sin\theta}^{+\sqrt{2AE}\,\sin\theta} \frac{dp_\varphi}{\left[2AE\sin^2\theta - p_\varphi^2\right]^{1/2}} \checkmark$$

Now: $$\int_{-a}^{a} \frac{dx}{(a^2-x^2)^{1/2}} = 2\int_0^a \frac{dx}{(a^2-x^2)^{1/2}}$$

$$= +2i \log\left(x + \sqrt{x^2-a^2}\right)\Big]_0^a = -\left(-2i\log a + 2i\log ia\right)$$

$$= +2i\log(-i) = -(+2i\log i) = -2i\log e^{i\pi/2}$$

$$= -2i\cdot i\pi/2 = \pi$$

$$\cos\theta = \frac{x}{a}, \quad dx = -a\sin\theta\, d\theta$$

$$a\sin\theta = \sqrt{a^2-x^2}$$

$$2\int_0^a \frac{dx}{(a^2-x^2)^{1/2}} = -2\int_{\pi/2}^{0} d\theta = \pi = \int_{-a}^{a} \frac{dx}{(a^2-x^2)^{1/2}}$$

Then:

$$\Omega(E) = 4\pi^2 A \int_0^{\pi} \sin\theta\, d\theta = -4\pi^2 A \int_1^{-1} d(\cos\theta)$$

$$= 8\pi^2 A$$

or: $\boxed{\Omega(E) = 8\pi^2 A}$ ✓

⑩

P262 : Problem 5

The spherical pendulum or the rigid rotator in an electric field.

$$\mathcal{H} = \frac{1}{2A}\left\{ p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} \right\} - \mathcal{E}d\,\cos\theta$$

d is the dipole moment, $-\mathcal{E}d < E$

Find $\Omega(E)$

We want to calculate:

$$\Omega(E) = \int \delta(E-\mathcal{H})\, dp_\varphi\, dp_\theta\, d\theta\, d\varphi$$

$$= 2\pi \int \delta(E-\mathcal{H})\, dp_\varphi\, dp_\theta\, d\theta$$

Define: $$f(p_\theta) = E - \frac{1}{2A}\left\{ p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} \right\} + \mathcal{E}d\,\cos\theta$$

$$\Omega(E) = 2\pi \int \delta\left\{ E - \frac{1}{2A}\left\{ p_\theta^2 + \frac{p_\varphi^3}{\sin^2\theta} \right\} + \mathcal{E}d\,\cos\theta \right\} dp_\theta\, dp_\varphi\, d\theta$$

Define: $x = \frac{p_\theta}{\sqrt{2AE}}$; $y = \frac{p_\varphi}{\sqrt{2AE}}$; $b = \frac{\mathcal{E}d}{E}$

$b \leq 1$

Then:

$$\Omega(E) = \frac{2\pi}{E}\cdot 2AE \int \delta\left(1 - x^2 - \frac{y^2}{\sin^2\theta} + b\cos\theta\right) dx\, dy\, d\theta$$

$$= 4\pi A \int \delta\left(1 - x^2 - \frac{y^2}{\sin^2\theta} + b\cos\theta\right) dx\, dy\, d\theta$$

Define: $f(x) = 1 - x^2 - \frac{y^2}{\sin^2\theta} + b\cos\theta$

$$x_0 = \sqrt{1 - \frac{y^2}{\sin^2\theta} + b\cos\theta} = \frac{1}{\sin\theta}\sqrt{(1+b\cos\theta)\sin^2\theta - y^2}$$

$$|f'(x_0)| = \frac{2}{\sin\theta}\sqrt{(1+b\cos\theta)\sin^2\theta - y^2}$$

Then:

$$\Omega(E) = 4\pi A \int_0^{\pi} d\theta \sin\theta \int_{-\sin\theta\sqrt{1+b\cos\theta}}^{\sin\theta\sqrt{1+b\cos\theta}} \frac{dy}{\left[(1+b\cos\theta)\sin^2\theta - y^2\right]^{1/2}}$$

$$= 8\pi^2 A$$

The limits on θ are governed by the roots of $(1+b\cos\theta)\sin^2\theta = 0$ whose only real roots are $\theta = 0, \pi, 2\pi, \ldots$ since $b \leq 1$ as stated in the problem.

Therefore $\boxed{\Omega(E) = 8\pi^2 A}$ for $E > |\mathcal{E}d|$

<u>See page 4:</u>

Another way to do this is to use:

$$\Omega(E) = \int \Omega_1(E_1)\,\Omega_2(E-E_1)\,dE_1$$

$$= 4\pi A \int \delta\left(E - x^2 - \frac{y^2}{\sin^2\theta} - b\cos\theta\right) dx\,dy\,d\theta \quad , \text{Here } b = -\frac{\mathcal{E}d}{E}$$

$$\Omega_1(E) = \int \delta(E - x^2)\,dx = \frac{1}{\sqrt{E}}$$

$$\Omega_2(E) = \int \delta\left(E - \frac{y^2}{\sin^2\theta} - b\cos\theta\right) dy\,d\theta$$

$$= \int_0^{\pi} \frac{\sin\theta\,d\theta}{\sqrt{E - b\cos\theta}} = \int_{-1}^{1} \frac{du}{\sqrt{E-bu}} \quad ; \quad \begin{aligned} w &= E - bu \\ dw &= -b\,du \end{aligned}$$

$$= \frac{1}{b}\int_{E-b}^{E+b} \frac{dw}{w^{1/2}} = \frac{2}{b}\left\{\sqrt{E+b} - \sqrt{E-b}\right\}$$

$$\Omega(E) = \frac{8\pi A}{b}\left\{\underbrace{\int_0^{E+b} \frac{\sqrt{E+b-E_1}}{\sqrt{E_1}}\,dE_1}_{\substack{2\int_0^{\sqrt{E+b}} \sqrt{E+b-s^2}\,ds \\ = (E+b)\,\pi/2}} - \underbrace{\int_0^{E-b} \frac{\sqrt{E-b-E_1}}{\sqrt{E_1}}\,dE_1}_{\substack{2\int_0^{\sqrt{E-b}} \sqrt{E-b-s^2}\,ds \\ = (E-b)\,\pi/2}}\right\}$$

or $$\boxed{\Omega(E) = 8\pi^2 A}$$

<u>See page 4:</u>

now consider the case $b = \frac{\mathcal{E}d}{E}$ but for which $-\mathcal{E}d < E$. $<\mathcal{E}d$ Hence $b \geq 1$ and we have a root given by:

$$1 - b\cos\theta = 0 \quad , \quad \cos\theta = \frac{1}{b} \quad , \quad \theta = \cos^{-1}\frac{1}{b}$$

Here $b = -\frac{\mathcal{E}d}{E}$, Thus we have for the structure function:

$$\Omega(E) = 4\pi A \int_0^{\cos^{-1}\frac{1}{b}} d\theta \sin\theta \int_{-\sin\theta\sqrt{1-b\cos\theta}}^{\sin\theta\sqrt{1-b\cos\theta}} \frac{dy}{\left[(1-b\cos\theta)\sin^2\theta - y^2\right]^{1/2}}$$

$$= 4\pi^2 A \int_{\cancel{0}\ 1/b}^{1} d(\cos\theta) = 4\pi^2 A \left[1 - \frac{1}{b}\right]$$

or:

$$\boxed{\Omega(E) = 4\pi^2 A \left[1 + \frac{E}{\mathcal{E}d}\right]}$$

And $\Omega(E) = 0$ for $E < -\mathcal{E}d$

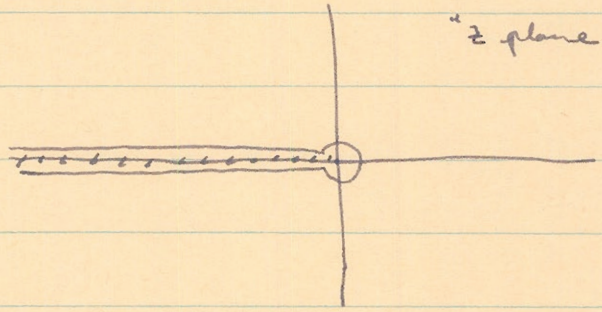
OK

⑩

P262 : Problem 6

Definition : $\frac{1}{\Gamma(n)} = \frac{1}{2\pi i} \int \frac{e^{z}}{z^{n}} \, dz$

Evaluate this integral by the saddle point method up to the term of order $\frac{1}{n}$, including the term of $O(\frac{1}{n})$. The contour of the above integral definition is :



Write the integral as : $\frac{1}{\Gamma(n)} = \frac{1}{2\pi i} \int e^{f(z)} \, dz$

where $f(z) = z - n \ln z$, n is real

We find the saddle point by taking $f'(z) = 0$.

$1 - \frac{n}{z} = 0$ or $z = n$ is the saddle point. We now must find the path of steepest descent which will be the path of stationary phase.

Im $f(z) =$ constant ;

$$f(z) = x + iy - n \ln(x+iy) \quad ; \quad (x+iy) = \sqrt{x^2+y^2}\, e^{i \tan^{-1}\frac{y}{x}}$$

$$f(z) = x + iy - \frac{n}{2} \ln(x^2+y^2) - i n \tan^{-1}\frac{y}{x}$$

$$\therefore \quad y - n \tan^{-1}\frac{y}{x} = \text{constant} = c$$

When $x = n$, $y = 0$, $\therefore$ $c = 0$, in the principle value.

Then:

$$y - n \tan^{-1}\frac{y}{x} = 0$$

is the path of steepest descent. Near the saddle point:

$$\frac{dy}{dx} - \frac{n}{1 + \left(\frac{y}{x}\right)^2}\left\{\frac{1}{x}\frac{dy}{dx} - \frac{y}{x^2}\right\} = 0$$

$$\tan^{-1}\frac{y}{x} = \frac{y}{n} \quad ; \quad x = \frac{y}{\tan\frac{y}{n}}$$

$$\frac{dx}{dy} = \frac{1}{\tan\frac{y}{n}} + y\frac{d}{dy}\left(\frac{1}{\tan\frac{y}{n}}\right)$$

$$\frac{d}{dy}\frac{1}{\tan\frac{y}{n}} = \frac{d}{dy}\cot\frac{y}{n} = -\frac{1}{n}\csc^2\frac{y}{n}$$

$$\therefore \quad \frac{dx}{dy} = \cot\frac{y}{n} - \frac{y}{n}\csc^2\frac{y}{n} = \cot\frac{y}{n} - \frac{y}{n\sin^2\frac{y}{n}}$$

$$= \frac{1}{\tan\frac{y}{n}} - \frac{y}{n\sin^2\frac{y}{n}}$$

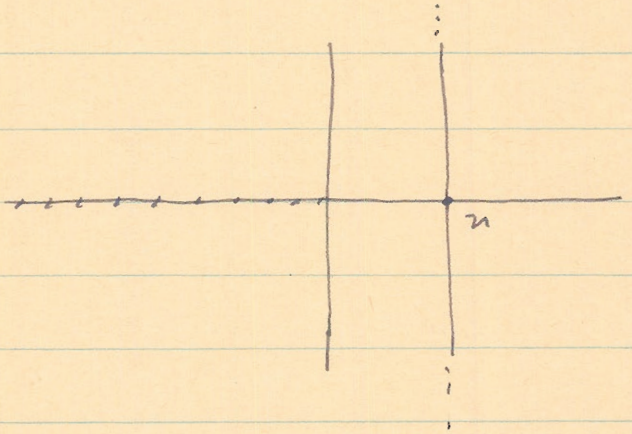
near $y=0$, $x=n$

$$\frac{dx}{dy} \cong \frac{1}{\frac{y}{n}} - \frac{y}{\frac{y^2}{n}} \cong \frac{1}{n}\left\{\frac{1}{y} - \frac{1}{y^{\cancel{2}}}\right\} = 0$$

as $y \to 0$

So at $x=n$, $y=0$, $\frac{dy}{dx} = \infty$

We deform the contour such that:



now:

$$\frac{1}{\Gamma(n)} = \frac{1}{2\pi i}\int_{n-i\infty}^{n+i\infty} \frac{e^z}{z^n}\,dz = \frac{e^n}{2\pi}\int_{-\infty}^{\infty} \frac{e^{iy}}{(n+iy)^n}\,dy$$

This is still exact.

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{f(n+iy)}\,dy$$

We now expand $f(z)$ about the saddle point:

$$f(z) = f(n) + f'(n)(z-n) + \frac{f''(n)(z-n)^2}{2} + \frac{f'''(n)(z-n)^3}{6} + \cdots$$

$$f(z) = z - n\ln z$$

$$f'(z) = 1 - \frac{n}{z}$$

$$f''(z) = \frac{n}{z^2} \quad ; \quad f'''(z) = \frac{-2n}{z^3} \quad ; \quad f^{iv}(z) = \frac{6n}{z^4}$$

Therefore:

$$f(z) = n - n\ln n + \frac{(z-n)^2}{2n} - \frac{(z-n)^3}{3n^2} + \frac{(z-n)^4}{4n^3}$$

$$f(n+iy) = n - n\ln n - \frac{y^2}{2n} + i\frac{y^3}{3n^2} + \frac{y^4}{4n^3}$$

Then:

$$\frac{1}{\Gamma(n)} \approx \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{n} n^{-n} e^{-\frac{y^2}{2n}} \left\{1 + i\frac{y^3}{3n^2} + \cdots\right\} dy$$

$$= \frac{e^{n} n^{-n}}{2\pi}\left\{\int_{-\infty}^{\infty} e^{-\frac{y^2}{2n}} dy + \frac{i}{3n^2}\int_{-\infty}^{\infty} y^3 e^{-\frac{y^2}{2n}} dy\right\}$$

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2n}} dy = \sqrt{2n\pi}$$

$$\int_{-\infty}^{\infty} y^3 e^{-\frac{y^2}{2n}} dy = 0$$

The next term in the expansion of $e^{iy^3/3n^2}$ is $i\,y^6$ so we have to go to the fourth term in the taylor series expansion: The integral is:

$$\frac{1}{4n^3}\int_{-\infty}^{\infty} y^4 e^{-\frac{y^2}{2n}} dy = \frac{1}{4n^3}\,2\cdot\frac{3}{8}\cdot 4n^2\sqrt{2\pi n}$$

$$= \frac{3}{4n}\sqrt{2\pi n}$$

the $i^2 y^6$ is just as important.

Therefore:

$$\frac{1}{\Gamma(n)} \simeq \frac{e^{n} n^{-(n+1/2)}}{\sqrt{2\pi}} \left\{ 1 + \frac{3}{4n} + \cdots \right\}$$

or $\quad O\left(\frac{1}{n}\right) = \frac{3}{4n}$ at the saddle point.

$$\log \Gamma(n) \simeq (n+1/2)\log n - n + \frac{1}{2}\log 2\pi - \log\frac{3}{4} + \log n + \cdots$$

⑥

Grant

47/60