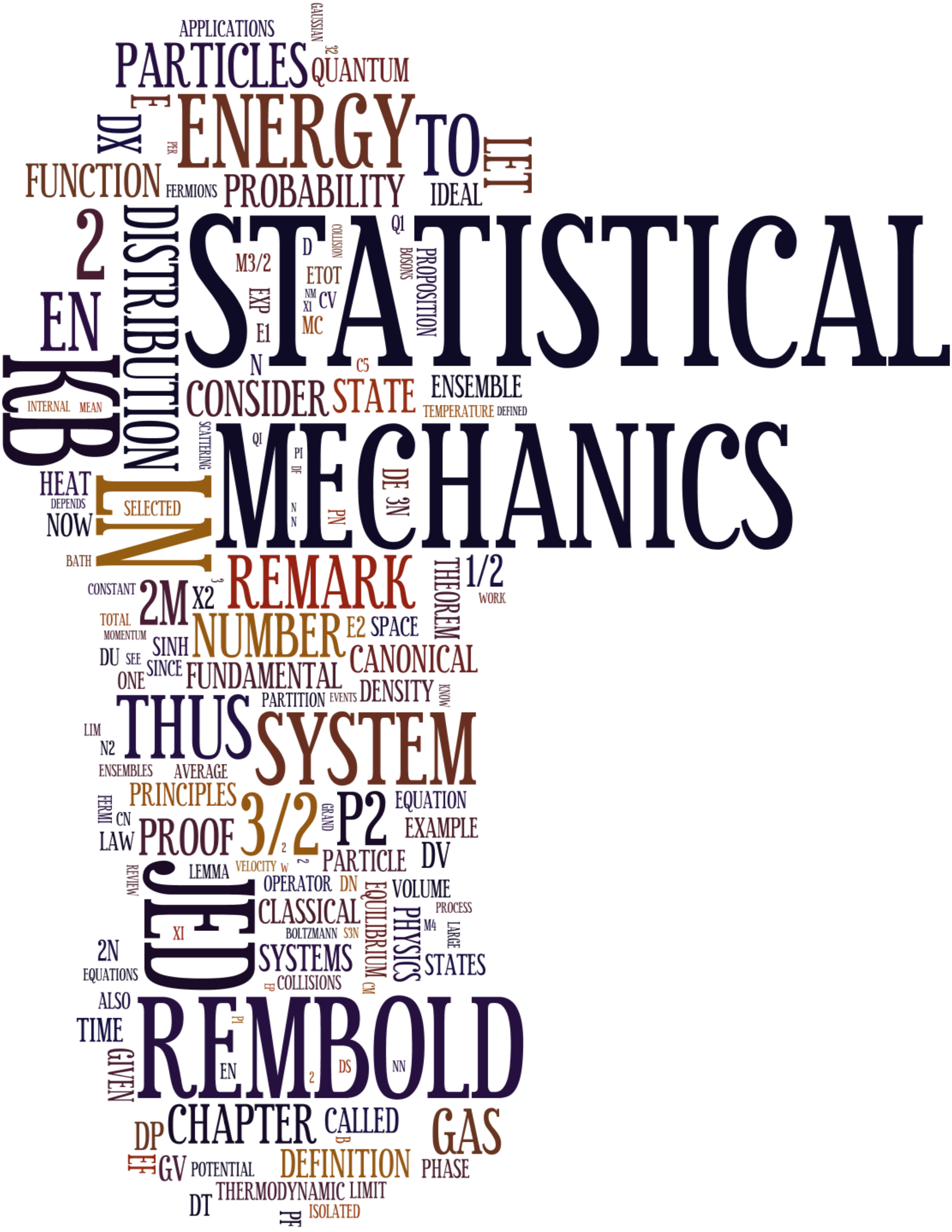

APPLICATIONS
GAUSSIAN
32
PARTICLES
QUANTUM
E
ENERGY
TO
LET
DX
PER
FUNCTION
FERMIONS
PROBABILITY
IDEAL
2
DISTRIBUTION
STATISTICAL
Q1
COLLISION
D
BOSONS
PROPOSITION
M3/2
ETOT
EXP
NM
X1
CV
MC
E1
EN
N
C5
KB
CONSIDER
STATE
ENSEMBLE
INTERNAL
MEAN
TEMPERATURE
DEFINED
SCATTERING
QI
PI
DF
MECHANICS
DE
3N
HEAT
DEPENDS
LN
SELECTED
N
N
PN
NOW
BATH
REMARK
THEOREM
1/2
WORK
CONSTANT
2M
X2
3
TOTAL
NUMBER
E2
SPACE
MOMENTUM
SINH
CANONICAL
DU
SEE
SINCE
FUNDAMENTAL
ONE
PARTITION
EVENTS
DENSITY
KNOW
LIM
THUS
SYSTEM
N2
ENSEMBLES
AVERAGE
PRINCIPLES
3/2
GRAND
P2
EQUATION
FERMI
CN
LAW
PROOF
EXAMPLE
2
2
PARTICLE
DV
REVIEW
LEMMA
VELOCITY
W
OPERATOR
DN
EQUILIBRIUM
VOLUME
JED
PROCESS
CLASSICAL
PHYSICS
XI
M4
LARGE
BOLTZMANN
S3N
2N
SYSTEMS
STATES
EQUATIONS
EP
COLLISIONS
CM
ALSO
P1
TIME
REMBOLD
GIVEN
EN
2
DS
NN
CHAPTER
CALLED
DP
B
GAS
EF
GV
POTENTIAL
DEFINITION
PHASE
THERMODYNAMIC
LIMIT
DT
PF
ISOLATED

# Statistical Mechanics Notes

Jed Rembold

September 10, 2014

# Acknowledgments

These notes are transcripts of lectures given by S.L. Sessions at New Mexico Tech during Spring 2011 and D. Belitz at the University of Oregon since 1989. They are based in part on notes taken by D. Belitz in lectures given by Wolfgang Götze at the Technical University of Munich between 1975 and 1985. The original handwritten notes were typeset by Jed Rembold.

# Contents

# Chapter 1

# Fundamental Principles of Statistical Physics

## 1.1 Review of Topics in Classical Mechanics

### 1.1.1 Hamilton's Equations and Phase Space

**Axiom**: The state of a system with $n$ degrees of freedom is completely determined by $n$ coordinates $q_i(t)$ and $n$ velocities $\dot{q}_i(t)$ (for $i = 1, \ldots, n$)

**Lagrangian**: The Lagrangian can be written as

$$\mathcal{L}(q, \dot{q}, t)$$

where

$$\dot{q}_i(t) = \frac{dq(t)}{dt} \quad \text{and} \quad q(t) = (q_1(t), q_2(t), \ldots, q_n(t))$$

**Momenta**: Similarly, the momenta $p_i(t)$ can be expressed as:

$$p_i(t) = \frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}_i}$$

**Definition 1:** The Hamiltonian or Legendre Transformation of $\mathcal{L}$ is given by the function:

$$H(q, p, t) := \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L}(q, \dot{q}, t)$$

**Definition 2:** The space of 2n-tuplets $(q, p) = (q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n)$ is called the phase space of the system, and is denoted by $\Gamma = \{(q, p)\}$.

**Theorem**: For every path $q(t)$ in $\mathbb{R}^n$, there is one and only one path $q(t), p(t)$ in $\Gamma$. The path in $\Gamma$ in determined by Hamilton's Equations or the Canonical Equations.

$$\dot{q}_i(t) = \frac{\partial H(q,p,t)}{\partial p_i}$$
$$\dot{p}_i(t) = -\frac{\partial H(q,p,t)}{\partial q_i}$$

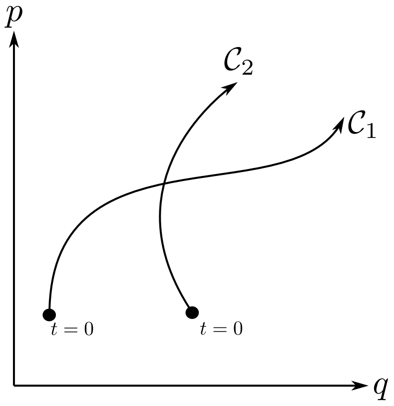
which are equivalent to the Euler-Lagrange equations:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i}$$

Proof: See Phys 505

***Remark 1:*** The Euler-Lagrange equations give $n$ 2nd order ODEs, which is equivalent to the $2n$ 1st order ODEs given by Hamilton's equations.

***Remark 2:*** Given an initial point $(q_0, p_0) \,\epsilon\, \Gamma$ with $q_0 = q(t=0)$, $p_0 = p(t=0)$ then this uniquely determines the path $q(t), p(t)$. Thus paths in phase space can not cross.



The above is *NOT* physically possible since $t = 0$ could be chosen to be at the intersection, which would give two different paths with identical initial points.

***Remark 3:*** Using chain rule:

$$\frac{d}{dt}H(q,p,t) = \sum_{i=1}^{n}\frac{\partial H}{\partial p_i}\dot{p}_i + \sum_{i=1}^{n}\frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial t}$$

Using Hamilton's equations,

$$\begin{aligned}&= \sum_{i=1}^{n} (\dot{q}_i \dot{p}_i - \dot{p}_i \dot{q}_i) + \frac{\partial H}{\partial t} \\ &= \frac{\partial H}{\partial t}\end{aligned}$$

So if the Hamiltonian has no explicit time dependence, then

$$\frac{d}{dt} H(q,p) = 0$$

which is a 1st integral (conservation law). Thus, as $H\left(q_0(t), p_0(t)\right) = E$ which is constant, energy is conserved or autonomous in the problem.

**Example**: Let's look at a 1-d Harmonic Oscillator

$$n = 1, \qquad \mathcal{L} = \tfrac{1}{2} m\dot{q}^2 - \tfrac{1}{2} kq^2, \qquad \Rightarrow p = \frac{\partial \mathcal{L}}{\partial \dot{q}} = m\dot{q}$$

$$\begin{aligned}H = p\dot{q} - \mathcal{L} &= m\dot{q}^2 - \tfrac{1}{2} m\dot{q}^2 + \tfrac{1}{2} kq^2 \\ &= \tfrac{1}{2} m\dot{q}^2 + \tfrac{1}{2} kq^2 \\ \Rightarrow H(q,p) &= \frac{p^2}{2m} + \frac{1}{2} kq^2\end{aligned}$$

The phase space orbit in then

$$E = H(q,p) = \frac{p^2}{2m} + \frac{1}{2} kq^2$$

$$\Rightarrow \frac{p^2}{2mE} + \frac{q^2}{2E/k} = 1$$

The canonical equations are:

$$\begin{aligned}\dot{q} &= \frac{\partial H}{\partial p} = \frac{p}{m} \\ \dot{p} &= -\frac{\partial H}{\partial q} = -kq \\ \Rightarrow \ddot{q} &= -\frac{k}{m} q = -\omega^2 q\end{aligned}$$

where $\omega = \sqrt{\frac{k}{m}}$ as per normal.

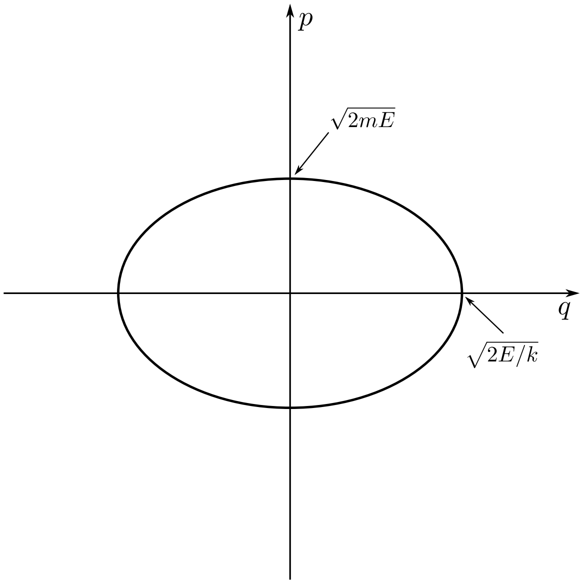


Figure 1.1: Phase space diagram of a simple harmonic oscillator with energy E.

### 1.1.2 Phase Flow

Let

$$\begin{aligned}(q(t), p(t)) &\equiv z(t) \in \Gamma \\ &= (q_1(t), q_2(t), \ldots, q_n(t), p_1(t), p_2(t), \ldots, p_n(t))\end{aligned}$$

**Definition 1:** The set of mapping $U : \Gamma \to \Gamma$,

$$U(t, t_0)z(t_0) = z(t)$$

is called phase flow or time-evolution.

**Proposition**: The time evolution operator, or propagator, $U(t, t_0)$, has the properties:

1. $U(t,t) = \mathbb{I} \quad \forall\, t$
2. $U(t, t_1)U(t_1, t_0) = U(t, t_0) \quad \forall\, t, t_0, t_1$
3. $U^{-1}(t, t_0) = U(t_0, t) \quad \forall\, t, t_0$

Proofs:

1.

$$U(t,t)z(t) = z(t) \quad \forall z(t) \in \Gamma$$

$$\therefore U(t,t) = \mathbb{I}$$

2.

$$\begin{aligned} U(t,t_1)U(t_1,t_0)z(t_0) &= U(t,t_1)z(t_1) = z(t) \\ &= U(t,t_0)z(t_0) \end{aligned}$$

$$\therefore U(t,t_1)U(t_1,t_0) = U(t,t_0)$$

3.

$$\begin{aligned} U(t_0,t)z(t) &= z(t_0) \\ U^{-1}(t_0,t)U(t_0,t)z(t) &= U^{-1}(t_0,t)z(t_0) \\ z(t) &= U^{-1}(t_0,t)z(t_0) \\ &= U(t,t_0)z(t_0) \end{aligned}$$

$$\therefore U^{-1}(t_0,t) = U(t,t_0)$$

***Remark 1:*** Knowledge of the phase flow is equivalent to knowledge of the solution of the canonical equations.

### 1.1.3 Liouville's Theorem

**Theorem**: Let

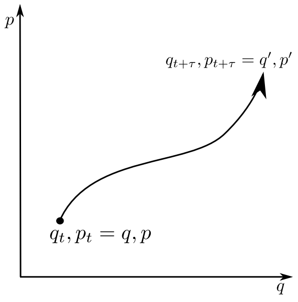
$$d\Gamma = \text{const} \cdot dq_1 dq_2 \cdots dq_n dp_1 dp_2 \cdots dp_n$$

be the element of volume in phase space. Then

$$d\Gamma = \text{ constant}$$

In other words, phase flow is volume preserving.

<u>Proof:</u>



Let $q, p = q_t, p_t$ represent $q, p$ at time $t$, and $q', p' = q_{t+\tau}, p_{t+\tau}$ represent values of $q, p$ at time $t + \tau$. Then we want to show that

$$\int dq_1 dq_2 \cdots dq_n dp_1 dp_2 \cdots dp_n = \int dq_1' d_2' \cdots dq_n' dp_1' dp_2' \cdots dp_n'$$

But

$$\int dq_1' \cdots dq_n' dp_1' \cdots dp_n' = \int J\, dq_1 \cdots dq_n dp_1 \cdots dp_n$$

where

$$J = \frac{\partial(q_1', \ldots, q_n', p_1', \ldots, p_n')}{\partial(q_1, \ldots, q_n, p_1, \ldots, p_n)}$$

$$= \begin{vmatrix} \frac{\partial q_1'}{\partial q_1} & \cdots & \frac{\partial q_1'}{\partial q_n} & \frac{\partial q_1'}{\partial p_n} & \cdots & \frac{\partial q_1'}{\partial p_n} \\ \vdots & & & & & \vdots \\ \frac{\partial q_n'}{\partial q_1} & & & & & \vdots \\ \frac{\partial p_1'}{\partial q_1} & & & & & \vdots \\ \vdots & & & & & \vdots \\ \frac{\partial p_n'}{\partial p_1} & \cdots & \cdots & \cdots & \cdots & \frac{\partial p_n'}{\partial p_n} \end{vmatrix}$$

where $J$ is the Jacobian and is the determinant of the above matrix. Proving Liouville's Theorem then boils down to showing that $J$=1. This can be shown using results from Advanced Dynamics:

1. The change in $p, q$ during motion may be regarded as a canonical transformation.
2. Volume in a phase space is invariant under canonical transformations (i.e. $J = 1$ for canonical transformations).

### 1.1.4 Poincare's Theorem

**Theorem**: *For full mathematical representation and proof, see "Statistical Mechanics" by Huang.* A system having a finite energy and confined to a finite volume will, after a sufficiently long time, return to an arbitrarily small neighborhood about almost any given initial state.

***Remark 1:*** Thus, $V$ returns to $U$ after a finite number of steps. This is called the Recurrence Theorem.

***Remark 2:*** Recurrence occurs arbitrarily often and arbitrarily accurately (since $U$ is arbitrary).

**Corollary**: Let's consider a conservative mechanical system at some initial time to be in a state given by a point in phase space, $z_0 \in \Gamma$. Then the system will come arbitrarily close to $z_0$ again in finite time.

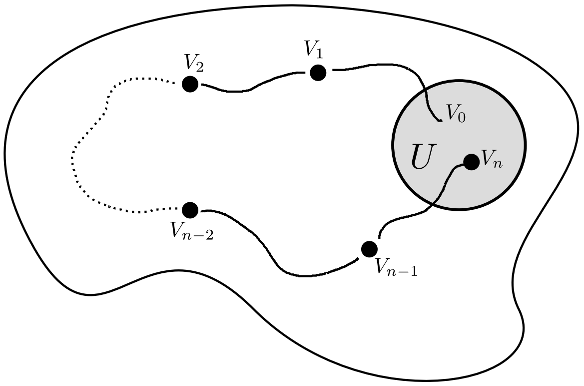
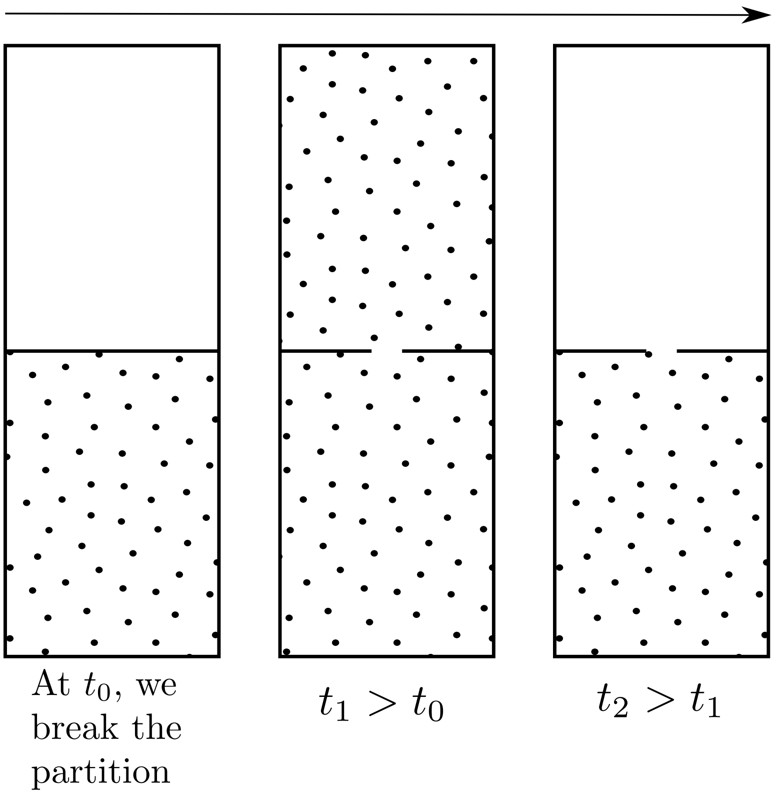


Figure 1.2: Graphical representation of Poincare's Theorem. The outer shape is some bounded region in $\mathbb{R}^2$ (for example), $V_0$ is the initial state of the system, $U$ is a neighborhood around $V_0$, and the lines mark the path of the system through time.

**Example**: Gas in a partitioned container



So, according to Poincare, at finite $t_2$, all the gas will go back into the single side of the container!

**Problem**: This is *NOT* what we observe!

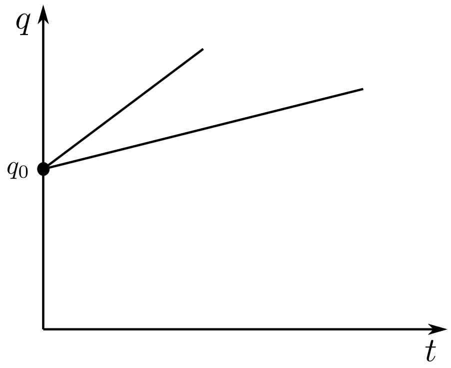
Solution:

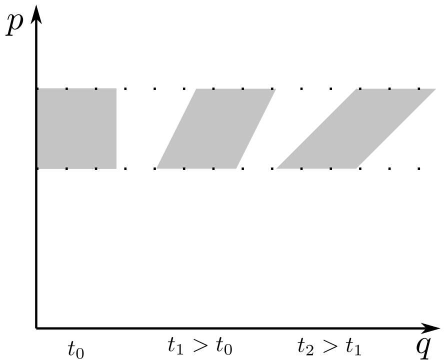
1. The “finite” time is ridiculously long for large systems.
2. Anything else? Yes! According to Boltzmann, this state would represent a “fluctuation from equilibrium and is thus still consistent with statistical description.

### 1.1.5 Necessity of a Statistical Description of Many Particle Systems

**Example 1**: Here we’ll look at a free particle in 1 dimension (which is thus not bounded).

$$p(t) = p_0$$
$$q(t) = q_0 + \left(\frac{p_0}{m}\right) t$$





By Liouville’s theorem, the volume in phase space in conserved. But it never returns to its initial state since the region is not bounded (and thus Poincare does *not* hold).

**Example 2**: Now we’ll look at a particle in a 1 dimensional potential well (so the system is now bounded).

$$p(t) = \pm p_0$$
$$q(t) = q_0 + \left(\frac{p_0}{m}\right) t + \text{ inverted motion}$$

As seen in the following figure, notice that $p$ is not continuous, and thus the region becomes disconnected. Thus we should modify the theory to allow for disconnected volumes. Also note that, as $t \to \infty$, the disconnected parts become thinner and thinner, which leads to spaghetti structure.

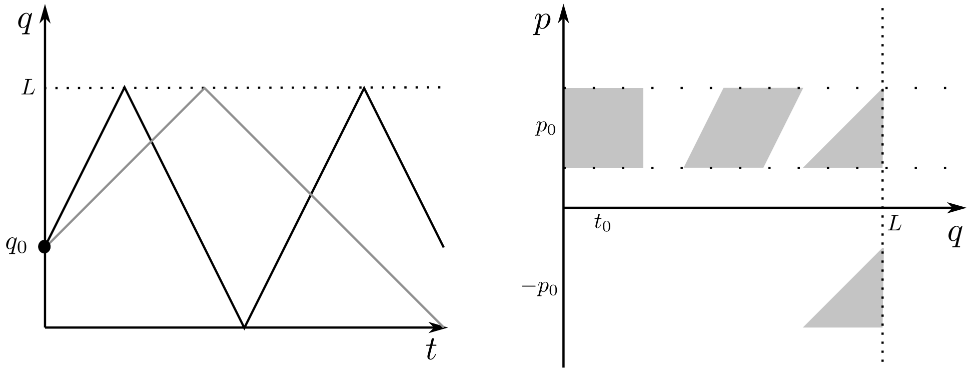
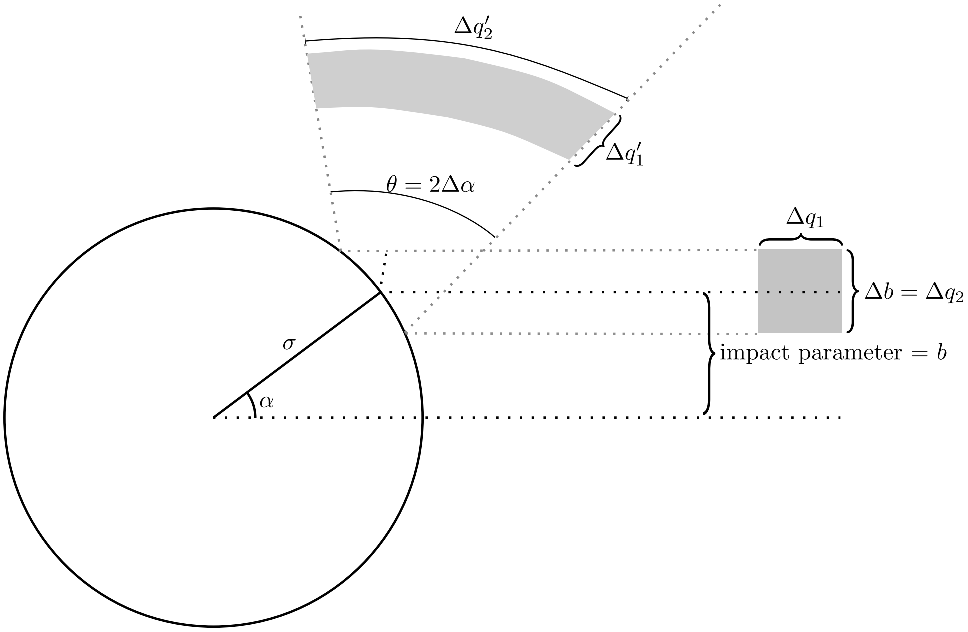


Figure 1.3: Spacial and phase space plots for a particle trapped in a potential well.

**Example 3**: For the final example, we'll look at a particle scattered by a hard sphere in 3 dimensions.



We know that

$$\sin\alpha = \frac{b}{\sigma}, \qquad \Delta q_1 = \Delta q'_1, \qquad \Delta q'_2 > \Delta q_2 = \Delta b$$

So the $q$ contribution to phase space volume increases due to scattering. This means that, in order to keep the phase space volume invariant, the moment contribution must decrease. Let

$$\alpha \ll 1 \quad \Rightarrow \quad \Delta\alpha = \frac{\Delta b}{\sigma} = \frac{\Delta q_2}{\sigma}$$

and let the mean-free path $l$ be

$$l \gg \sigma.$$

Therefore, the new spacial volume after 1 collision would be:

$$\begin{aligned} V' &= \Delta q_1' \cdot l \cdot 2\Delta\alpha \qquad \text{(in 2D)} \\ &= \Delta q_1' \cdot 2l \cdot \frac{\Delta q_2}{\sigma} \\ &= V \cdot \frac{2l}{\sigma} \end{aligned}$$

where $V =$ the initial volume $= \Delta q_1 \Delta q_2$. In $d$-dimensions then:

$$V' \sim V \left(\frac{2l}{\sigma}\right)^{d-1}$$

Now Louiville's theorem tells us that the momentum-space contribution to the phase space volume must decrease by a factor of $\left(\frac{\sigma}{2l}\right)^{(d-1)N}$ after $N$ collisions.

**Problem - "The Butterfly Effect"**: After $N$ collisions, a simple volume in phase space has been transformed into spaghetti in momentum space with typical dimensions given by

$$\frac{\Delta p_N}{p_N} \sim \left(\frac{\sigma}{2l}\right)^{(d-1)N}$$

**Example**: Consider a classical gas of hard spheres of radius $\sigma$ and number density $n$.

$$l \sim (n\sigma^2)^{-1}$$

$$\begin{aligned} \Rightarrow \frac{\sigma}{l} &\sim n\sigma^3 \equiv n^\star \quad \leftarrow \text{ dimensionless density} \\ &\approx 10^{-3} \text{ for typical gas} \end{aligned}$$

For 3 dimensions and 3 collisions ($D = 3, N = 3$):

$$\frac{\Delta p_N}{p_N} \sim (10^{-3})^{2\cdot 3} = 10^{-18}$$

So we would have to keep this accuracy to track the particles accurately. Thus, after a few collisions, a ridiculous level of accuracy is needed to maintain the spaghetti structure.

***Remark 1:*** Consider a real experiment, where we will be looking at the effect of gravitational attraction between an external butterfly and our system.

$$\frac{F}{m} = \frac{GM}{r^2} = 7 \times 10^{-11} \frac{M}{r^2}$$

The momentum uncertainty due to $F$ is:

$$\Delta p = F\tau$$

where $\tau$ is the mean free time between particles.

$$\tau = \frac{l}{v}$$

$$\Rightarrow \frac{\Delta p}{p} \sim \left(\frac{F\tau}{mv}\right) = \left(\frac{F}{m}\right)\left(\frac{\tau}{v}\right)$$

For a classical gas: $v = 10^3$ m/s, $\tau \sim 10^{-9}$ s.

$$\Rightarrow \frac{\Delta p}{p} \sim 10^{-22}\frac{M}{r^2}$$

- If the butterfly is $m = 1$ gram and $r = 100$ m.

  $$\frac{\Delta p}{p} \sim 10^{-22-7} = 10^{-29}$$

  We know spaghetti structure is destroyed after

  $$\frac{\Delta p_N}{p_N} \sim \left(\frac{\sigma}{2l}\right)^{(d-1)N}$$

  Setting $\frac{\Delta p}{p} = \frac{\Delta p_N}{p_N}$ and solving for $N$ yields:

  $$N = 5$$

  Thus, after only 5 collisions, the butterfly outside will affect the spaghetti structure of the system inside.

- If instead the butterfly (still $m = 1$ gram) is on the star Sirius ($r = 8.5$ light-years or $8 \times 10^{16}$ m):

  $$\frac{\Delta p}{p} = 10^{-22-35} = 10^{-57}$$

  $$\Rightarrow N = 9 \text{ collisions (which takes about } 100\mu\text{s}$$

  So again, we are screwed up even by butterflies on Sirius! Now, in order for the "picture" to reappear in the recurrence relation, the spaghetti structure must be conserved. Thus, uncontrollable forces in the microscopic world destroy spaghetti structure.

***Remark 2:*** If, instead, we were to consider a computer experiment, we would be free of outside effects, but would instead introduce uncontrollable perturbations in the form of rounding errors.

**Conclusions**:

1. Microscopically deterministic descriptions of N-particle systems over macroscopic times (with time reversibility, Poincare's theorem, etc) are impossible due to uncontrollable weak forces.
2. In experience, time average of macroscopic quantities are independent of the instabilities in phase-space structure.
3. We must develop a statistical, rather than deterministic, description of many particle systems.

***Remark 3:*** According to Boltzmann, the recurrence of states are fluctuations, which are certain to occur if you wait long enough. Thus, recurrence relations are completely consistent with a statistical viewpoint.

## 1.2 Statistical Methods

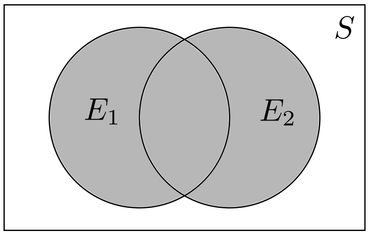
### 1.2.1 Probability

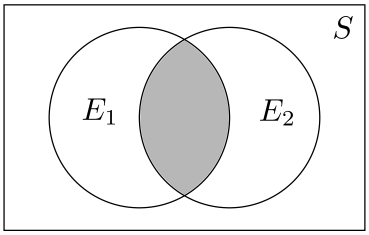
**Definition 1:** Consider a set $S \neq \emptyset$ which is the set of all possible outcomes of an experiment. This is called the sample set.

**Definition 2:** Let $E$ be any subset of $S$. $E$ is called an event and satisfies:

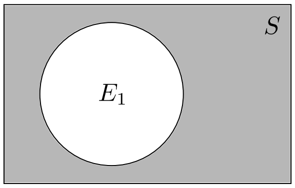
1. $E \subset S \; \forall E$
2. $\forall \, E_1, E_2 \subset S$, $(E_1 \cup E_2) \subset S \;\wedge\; (E_1 \cap E_2) \subset S$

   "For events $E_1$ and $E_2$ that are both possible, then $E_1$ or $E_2$ must also be possible, as well as $E_1$ and $E_2$ being possible."





3. $E \subset S, \; S/E \subset S$



**Definition 3:** Each event is assigned a statistical weight or probability $P(E)$ so that:

1. $0 \leq P(E) \leq 1 \; \forall E$
2. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
3. $P(S) = 1, \quad P(\emptyset) = 0$

**Example**: We will consider rolling a fair die:

$$S = \{1, 2, 3, 4, 5, 6\}$$

$$\text{Any } E \subset S \text{ is an event} \begin{cases} E = 1, \\ E = \{1, 3, 5\}, \\ \text{etc} \end{cases}$$

$$P(E) = \frac{(\text{number of elements in } E)}{6}$$

**Proposition**: The probability of mutually exclusive events equals the sum of the individual probabilities.

Proof: If you have mutually exclusive events, then $E_1 \cap E_2 = \emptyset$. Thus

$$\begin{aligned} P(E_1 \cup E_2) &= P(E_1) + P(E_2) - \underbrace{P(E_1 \cap E_2)}_{0} \\ &= P(E_1) + P(E_2) \end{aligned}$$

**Definition 4:** Let $E_1$ and $E_2$ be events. The joint probability of $E_1$ and $E_2$ is defined as

$$P(E_1, E_2) = P(E_1 \cap E_2)$$

**Definition 5:** If $E_1$ and $E_2$ are statistically independent, then

$$P(E_1, E_2) = P(E_1)P(E_2)$$

**Example**: Probability of getting either a 1 or a 6 in rolling 1 die once? The events are mutually exclusive, so

$$P(1 \cup 6) = P(1) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

**Example**: Probability of getting a 1 *and* a 6 when rolling 2 die once?

$$P(1, 6) + P(6, 1) = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{18}$$

Also

$$\begin{aligned} S = \{&(1,1),(1,2),\ldots,(1,6),\\ &(2,1),(2,2),\ldots\\ &\vdots\\ &(6,1),(6,2),\ldots,(6,6)\} \end{aligned}$$

All possible events are equally likely. Odds are 2 out of 36 possible

$$\frac{2}{36} = \frac{1}{18}$$

***Remark 1:*** It is non-trivial to determine whether events are statistically independent!

**Definition 6:** A probability distribution is the set of probabilities $P_i$ so that

$$\sum_i P_i = 1$$

A quantity given by a probability distribution is called a random variable.

### 1.2.2 Binomial Distribution

**Theorem**: Let there be $S$ independent attempts, and let the probability of a certain event in each attempt be $p$. The probability of an event occurring $n$ times is:

$$\rho_b(n,s) = \binom{S}{n} p^n (1-p)^{S-n} \qquad (\star)$$

where

$$\binom{S}{n} = \frac{S!}{n!(S-n)!}$$

Proof: The number of ways to partition $n$ events and $(S-n)$ non-events (out of $S$ total attempts) is

$$\binom{S}{n}$$

And so

$$\rho_b(n,) = \binom{S}{n}\rho^*(n,S)$$

where $\rho^*$ is the probability of realizing a particular partition. The probability for an event is $p$; thus the probability for a non-event is $(1-p)$. Thus

$$\rho^* = p^n(1-p)^{S-n}$$

***Remark 1:*** $\rho_b$ is called the binomial distribution.

***Remark 2:***

$$\sum_{n=0}^{S} \rho_b(n,S) = \sum_{n=0}^{S} \binom{S}{n} p^n (1-p)^{S-n}$$

Recall from the binomial theorem that

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Thus

$$\begin{aligned} \sum_{n=0}^{S} \rho_b(n,S) &= (p+1-p)^S \\ &= 1^S \\ &= 1 \end{aligned}$$

***Remark 3:*** Looking at an equivalent type problem, consider sites or boxes occupied with either 0 or 1 indistiguisable objects. Picking $S$ sites, what is the probability of finding $n$ objects on the $S$ sites, where on average $n = m$?

$$\begin{aligned} \rho_b(n,S;m) &= \binom{S}{n} \left(\frac{m}{S}\right)^n \left(1-\frac{m}{S}\right)^{S-n} \\ &= \frac{S!}{n!(S-n)!} \frac{m^n (S-m)^{S-n}}{S^S} \end{aligned}$$

### 1.2.3 Poisson Distribution

**Proposition**: A continuum of positions can be occupied by indistinguishable objects. Let the average number of objects in any sub-region by $m$. Then the probability of finding $n$ objects in that sub-region is:

$$\rho_p(n;m) = \frac{1}{n!} m^n e^{-m}$$

This is the Poisson Distribution.

**Example**: Consider a Geiger counter subject to constant radio activity. The probability of $n$ clicks/sec if, on average, there are $m$ clicks/sec, is given by a Poisson distribution. (Let the number of boxes go to infinity while keeping the density constant.)

**Lemma - Stirlings Formula**: For $n >> 1$,

$$n! = \sqrt{2\pi n}\, n^n e^{-n} \left[1 + \mathcal{O}(\tfrac{1}{n})\right]$$

This can also be written:

$$\ln n! = n \ln n - n + \tfrac{1}{2}\ln(2\pi n) + \mathcal{O}(\tfrac{1}{n})$$

<ins>Proof:</ins> See Reif Appendix A6

**Proof of Proposition**: By §1.2.2.R3,

$$\rho_b(n, S; m) = \frac{S!}{n!(S-n)!}\frac{m^n(S-m)^{S-n}}{S^S}$$

Let $S \to \infty$ while keeping $n, m$ fixed. Then

$$\rho_p = \rho_b(n, S \to \infty; m)$$

Consider

$$\begin{aligned}
\ln\left[\frac{n!\rho_b}{m^n}\right] &= \lim_{S\to\infty} \ln\left[\frac{S!}{(S-n)!}\frac{(S-m)^{S-n}}{S^S}\right] \\
&= \lim_{S\to\infty} \left[\ln S! - \ln(S-n)! + (S-n)\ln(S-m) - S\ln S\right] \\
&= \lim_{S\to\infty} \big[S\ln S - S + \tfrac{1}{2}\ln(2\pi S) + \mathcal{O}(\tfrac{1}{S}) - (S-n)\ln(S-n) \\
&\qquad + \tfrac{1}{2}\ln(2\pi(S-n)) + \mathcal{O}(\tfrac{1}{S-n} + (S-n)\ln(S-m) - S\ln S\big] \\
&= \lim_{S\to\infty} \big[\tfrac{1}{2}\ln(2\pi) + \tfrac{1}{2}\ln(S) - S\ln S - S\ln(1-\tfrac{n}{S}) + n\ln S + n\ln(1-\tfrac{n}{S}) \\
&\qquad - n - \tfrac{1}{2}\ln(2\pi) - \tfrac{1}{2}\ln S - \tfrac{1}{2}\ln(1-\tfrac{n}{S}) + S\ln S + S\ln(1-\tfrac{m}{S}) \\
&\qquad - n\ln S - n\ln(1-\tfrac{m}{S}) + \mathcal{O}(\tfrac{1}{S})\big] \\
&= \lim_{S\to\infty} \left[n - n - m + \mathcal{O}(\tfrac{1}{S})\right] \\
&= -m
\end{aligned}$$

$$\Rightarrow \rho_p = \frac{1}{n!}m^n e^{-m}$$

### 1.2.4 Averages and Fluctuations

Let $\rho(n)$ be a probability distribution for some event to occur $n$ times. Let $f(n)$ be some function of $n$.

**Definition 1:**

$$\langle f(n)\rangle_\rho := \sum_n f(n)\rho(n)$$

is called the average or mean of $f(n)$.

***Remark 1:*** $\langle f(n)\rangle$ is independent of $n$.

***Remark 2:*** Averages are defined with respect to a particular distribution.

***Remark 3:*** Choose $f(n) = n^m$:

$$\langle n^m\rangle = \sum_n n^m \rho(n)$$

This is called the m$^{th}$ moment of the distribution of $\rho(n)$.

***Remark 4:*** The zeroth moment is

$$\left\langle n^0\right\rangle = \langle 1\rangle = \sum_n \rho(n) = 1 \quad \text{by normalization}$$

**Example**: The first moment of $\rho_p(n;m)$:

$$\begin{aligned}\langle n\rangle_{\rho_p} &= \sum_{n=0}^{\infty} n\frac{1}{n!}m^n e^{-m}\\ &= \sum_{n=1}^{\infty} \frac{1}{(n-1)!}m^n e^{-m}\\ &= \sum_{n=0}^{\infty} \frac{1}{n!}m^{n+1}m^{-m}\\ &= me^{-m}\underbrace{\sum_{n=0}^{\infty}\frac{1}{n!}m^n}_{e^m}\\ &= m\end{aligned}$$

which agrees with the definition of the Poisson distribution! Yay!

**Definition 2:**

- $\left\langle(\Delta n)^2\right\rangle_\rho = \left\langle(n-\langle n\rangle)^2\right\rangle_\rho$ is called the mean squared deviation, the variance, or the 2nd moment abouth the mean of the distribution.

- $(\Delta n) := \left\langle (\Delta n)^2 \right\rangle^{1/2}$ is called the root mean squared (RMS) deviation or the standard deviation. Common notation has $\sigma = \Delta n$ and it gives us a measure of the width of the distribution.
- $\frac{\Delta n}{\langle n \rangle}$ is called the relative fluctuation.

**Example**: The 2nd moment of $\rho_p(n;m)$ is:

$$\begin{aligned}
\left\langle n^2 \right\rangle_{\rho_p} &= \sum_{n=0}^{\infty} n^2 \frac{1}{n!} m^n e^{-m} \\
&= m e^{-m} \sum_{n=0}^{\infty} \frac{(n+1)}{n!} m^n \\
&= m e^{-m} \left( m e^m + e^m \right) \\
&= m^2 + m
\end{aligned}$$

$$\begin{aligned}
\Delta n = \left\langle (n - \langle n \rangle)^2 \right\rangle_{\rho_p}^{1/2} &= \left( \left\langle n^2 \right\rangle - \langle n \rangle^2 \right)^{1/2} \\
&= \left( m^2 + m - m^2 \right) \\
&= m^{1/2}
\end{aligned}$$

***Remark 5:*** For large numbers of events, the size of the expected deviations from the mean, measured in units from the mean, goes to 0. In other words, the distribution becomes sharply peaked. This is known as the Law of Large Numbers.

### 1.2.5 Continuous Distributions

**Definition 1:** Let $x$ be a continuous random variable, and let $\rho(x)$ be the probability density that the variable has the value $x$. Then $\rho(x)\,dx$ is the probability that the value will lie in the interval between $x$ and $x + dx$. Let $\int \rho(x)\,dx = 1$. Then $\rho(x)$ is the (continuous) probability distribution function.

**Definition 2:** Moments, etc are defined by generalizations of §1.2.4.

$$\langle x^n \rangle = \int x^n \rho(x)\,dx \qquad \text{nth moment}$$

$$\sigma \equiv \Delta x := \left( \left\langle x^2 \right\rangle - \langle x \rangle^2 \right)^{1/2} \qquad \text{RMS Deviation}$$

**Example**: A factory produces resistors with a nominal resisitivity of $R$. Actual resistivities will be distributed according to some distribution function $\rho(R)$.

***Remark 1:*** We can always return to discrete values of the random variable by considering $\delta$-function distributions.

**Example**: Suppose $x$ can take on only integer values between 1 and $N$ with equal probability.

$$\Rightarrow \rho(x) = \frac{1}{N}\sum_{n=1}^{N}\delta(x-n)$$

is a continuous distribution function describing the above situation.

$$\int \rho(x)\,dx = 1$$

$$\begin{aligned}\int x\rho(x)\,dx &= \frac{1}{N}\sum_{n=1}^{N}\int x\delta(x-n)\,dx \\ &= \frac{1}{N}\sum_{n=1}^{N} n \\ &= \frac{1}{N}\left(\tfrac{1}{2}N(N+1)\right) \\ &= \tfrac{1}{2}(N+1)\end{aligned}$$

### 1.2.6 The Gaussian Distribution and the Central Limit Theorem

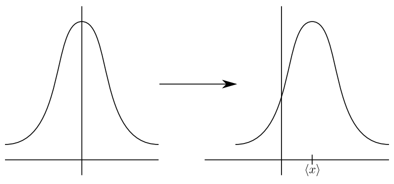
**Definition 1:** Let $x$ be a random variable with distribution

$$\rho_G(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2}$$

Then $\rho_G(x)$ is called a Gaussian or normal distribution.

***Remark 1:*** $\langle x\rangle = 0$. This can be trivially generalized by making the following substitution

$$x \to x - \langle x\rangle$$



***Remark 2:*** $\sigma$ is the standard deviation of $\rho_G$:

$$\left\langle x^2 \right\rangle = \frac{1}{\sigma\sqrt{2\pi}} \int x^2 e^{-x^2/2\sigma^2}\, dx$$

Scaling $x \to \sqrt{2}\sigma x$

$$= \frac{1}{\sigma\sqrt{2\pi}} 2\sigma^2\sqrt{2}\sigma \underbrace{\int x^2 e^{-x^2}\, dx}_{\frac{1}{2}\sqrt{\pi}}$$

$$= \sigma^2$$

**Lemma**: The moments of the Gaussian distribution are:

$$\left\langle x^n \right\rangle = \begin{cases} 0 & \text{if n = even} \\ \frac{n!\sigma^n}{2^{n/2}\left(\frac{n}{2}\right)!} & \text{if n = odd} \end{cases}$$

<u>Proof:</u>

$$\left\langle x^{2n+1} \right\rangle = 0 \qquad \text{by symmetry}$$

$$\begin{aligned} \left\langle x^{2n} \right\rangle &= \frac{1}{\sigma\sqrt{1\pi}} \int_{-\infty}^{\infty} x^{2n} e^{-x^2/2\sigma^2}\, dx \quad x \to \sqrt{2}\sigma x \\ &= \frac{1}{\sigma\sqrt{2\pi}} \left(\sqrt{2}\sigma\right)^{2n+1} 2 \int_0^\infty x^{2n} e^{-x^2}\, dx \end{aligned}$$

Now, looking at

$$\int_0^\infty x^{2n} e^{-x^2}\, dx$$

recall that

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

Letting $t = x^2$, we see that

$$\begin{aligned} \Rightarrow \int_0^\infty x^{2n} e^{-x^2}\, dx &= \int_0^\infty \tfrac{1}{2} t^{n-1} e^{-t}\, dt \\ &= \int_0^\infty \tfrac{1}{2} t^{n+\frac{1}{2}-1} e^{-t}\, dt \\ &= \tfrac{1}{2}\Gamma\left(n+\tfrac{1}{2}\right) \end{aligned}$$

$$\begin{aligned} \Rightarrow \left\langle x^{2n} \right\rangle &= \frac{1}{\sigma\sqrt{2\pi}} \left(\sqrt{2}\sigma\right)^{2n+1} \cdot 2 \cdot \tfrac{1}{2}\Gamma\left(n+\tfrac{1}{2}\right) \\ &= \frac{1}{\sqrt{\pi}} \left(2\sigma^2\right)^n \Gamma(n+\tfrac{1}{2}) \end{aligned}$$

Since $\Gamma(z)\Gamma(z+\frac{1}{2}) = \frac{\sqrt{\pi}\Gamma(2z)}{2^{2z-1}}$

$$\begin{aligned}
&= \frac{(2\sigma^2)^n}{\sqrt{\pi}} \frac{\sqrt{\pi}\Gamma(2n)}{\Gamma(n)2^{2n-1}} \\
&= \frac{2^n\sigma^{2n}(2n-1)!}{2^{2n-1}(n-1)!} \\
&= 2\left(\frac{2}{4}\right)^n \sigma^{2n} \frac{(2n)!}{2n}\frac{n}{n!} \\
&= \frac{\sigma^{2n}(2n)!}{2^n n!}
\end{aligned}$$

**Definition 2:** The characteristic function $f(x)$ of a distribution $\rho(x)$ is defined as:

$$f(k) = \left\langle e^{-ikx} \right\rangle_\rho = \int e^{-ikx}\rho(x)\,dx$$

(ie. the fourier transform of $\rho(x)$)

**Lemma**: For a gaussian distribution

$$f(k) = e^{-k^2\sigma^2/2}, \qquad \rho(x) = \int \frac{1}{2\pi} f(k) e^{ikx}\,dk$$

(since the fourier transform of a gaussian is a gaussian).

Proof:

$$\begin{aligned}
f(k) &= \left\langle e^{-ikx} \right\rangle_\rho \\
&= \sum_{n=0}^{\infty} \left\langle \frac{(-ikx)^n}{n!} \right\rangle_\rho \\
&= \sum_{n=0}^{\infty} \left\langle \frac{(-ikx)^{2n}}{(2n)!} \right\rangle_\rho \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n k^{2n}}{(2n)!} \left\langle x^{2n} \right\rangle \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n k^{2n}}{(2n)!} \frac{(2n)!\sigma^{2n}}{2^n n!} \\
&= \sum_{n=0}^{\infty} \frac{\left(\frac{-k^2\sigma^2}{2}\right)^n}{n!} \\
&= e^{-k^2\sigma^2/2}
\end{aligned}$$

$$\begin{aligned}
\int \frac{1}{2\pi} f(k) e^{ikx}\, dk &= \int e^{-k^2\sigma^2/2} e^{ikx} \frac{dk}{2\pi} \\
&= \frac{1}{2\pi} \int \exp\left[\frac{-\sigma^2}{2}\left(k^2 - \frac{2i}{\sigma^2}kx - \frac{x^2}{\sigma^4} + \frac{x^2}{\sigma^4}\right)\right] dk \\
&= \frac{1}{2\pi} e^{-x^2/2\sigma^2} \int \exp\left[-\frac{\sigma^2}{2}\left(k - \frac{ix}{\sigma^2}\right)^2\right] dk \\
&= e^{-x^2/2\sigma^2} \int \exp\left(-\frac{\sigma^2}{2}k\right) \frac{dk}{2\pi} \\
&= e^{-k^2/2\sigma^2} \frac{1}{2\pi} \frac{\sqrt{2}}{\sigma} \int e^{-k^2}\, dk \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} e^{-k^2/2\sigma^2} \\
&= \rho(k)_G
\end{aligned}$$

**Theorem: Central Limit Theorem**: Let $x_i$ $(i = 1, \ldots, N)$ be $N$ independent random variables with $\langle x_i \rangle = 0 \quad \forall i$ and $\left\langle x_i^2 \right\rangle < \infty \quad \forall i$. Also let $y = \frac{1}{N}\sum_{i=1}^{\infty} x_i$ be the arithmetic mean (ie. the average with respect to a flat distribution) of $x_i$. Then, for $N \to \infty$, $y$ is gaussian distributed with a standard deviation of $\sigma \sim \frac{1}{\sqrt{N}}$.

$$\rho(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-i^2/2\sigma^2}$$

with

$$\sigma^2 = \frac{1}{N^2} \sum_{i=1}^{N} \left\langle x_i^2 \right\rangle$$

***Remark 3:*** According to the law of large numbers, for large numbers of events the distribution is sharply peaked about its mean. And the CLT (Central Limit Theorem) tells us that the distribution in that limit is gaussian.

***Remark 4:*** This holds irrespective of what the distribution of $x$ is, so long as $\left\langle x_i^2 \right\rangle < \infty$.

***Remark 5:*** Generalizing to $\langle x_i \rangle \neq 0$:

$$\rho(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(y - \langle y \rangle\right)^2)/2\sigma^2\right]$$

with

$$\sigma^2 = \frac{1}{N^2} \sum_{i=1}^{N} \left\langle (x_i - \langle x_i \rangle)^2) \right\rangle$$

**Simple Proof of Theorem**:

$$\begin{aligned}
\left\langle e^{-iky}\right\rangle &= \left\langle \exp\left(-ik\frac{1}{N}\sum_{i=1}^{N} x_i\right)\right\rangle \\
&= \left\langle \prod_{i=1}^{N} e^{-ikx_i/N}\right\rangle \\
&= \prod_{i=1}^{N}\left\langle e^{-ikx_i/N}\right\rangle \qquad \text{since independent } x_i\text{'s} \\
&= \exp\left[\ln\left(\prod_{i=1}^{N}\left\langle e^{-ikx_i/N}\right\rangle\right)\right] \\
&= \exp\left[\sum_{i=1}^{N}\ln\left\langle e^{-ikx_i/N}\right\rangle\right] \\
&:= \exp\left[\sum_{i=1}^{N} A_i\left(\frac{k}{N}\right)\right]
\end{aligned}$$

where

$$A_i(p) = \ln\left\langle e^{-ipk_i}\right\rangle$$

Now, for $N \to \infty$, we expand:

$$\begin{aligned}
A_i\left(\frac{k}{N}\right) &= \ln\left\langle 1 + \frac{ik}{N}x_i - \tfrac{1}{2}\left(\frac{k}{N}\right)^2 x_i^2\right\rangle \\
&= \ln\left[1 - \tfrac{1}{2}\left(\frac{k}{N}\right)^2 \langle x_i\rangle^2 + \cdots\right] \\
&= \frac{-k^2}{2N^2}\left\langle k_i^2\right\rangle + \mathcal{O}(\tfrac{1}{N^3})
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \left\langle e^{-iky}\right\rangle &= \exp\left[-\sum_{i=1}^{N}\frac{k^2}{2N^2}\langle x_i\rangle^2\left[1+\mathcal{O}(\tfrac{1}{N})\right]\right] \\
&= e^{-k^2\sigma^2/2}
\end{aligned}$$

where

$$\sigma^2 = \frac{1}{N^2}\sum_{i=1}^{N}\left\langle x_i^2\right\rangle$$

By the Lemma then, $\rho(y)$ is gaussian where:

$$\rho(y) = \frac{\sigma\sqrt{2\pi}}{e^{-y^2/2\sigma^2}} \qquad \text{with} \qquad \sigma^2 = \frac{1}{N^2}\sum_{i=1}^{N}\left\langle x_i^2\right\rangle$$

Finally, let

$$\chi = \sup\left\{\left\langle x_i^2\right\rangle\right\}$$

Then

$$\sigma^2 \leq \frac{1}{N^2}\chi N = \frac{\chi}{N}$$
$$\Rightarrow \sigma \sim \frac{1}{\sqrt{N}}$$

**Example**: By §2.2, we have a binomial situation of $S$ total events with:

$$n_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } (1-p) \end{cases}$$
$$\Rightarrow \rho(n_i) = p\delta(n_i - 1) + (1-p)\delta(n_i)$$

Consider

$$n = \sum_{i=1}^{S} n_i$$

so that $n$ obeys a binomial distribution. Then

$$\rho(n) = \rho_b(n; s) = \binom{S}{n} p^N (1-p)^{S-n}$$

$$\langle n_i \rangle = \frac{1}{S}\sum_{i=1}^{S} n_i = p = \frac{n}{S}$$

$$\left\langle n_i^2 \right\rangle = \frac{1}{S}\sum_{i=1}^{S} n_i^2 = p \qquad \text{since } n_i = \{0, 1\} \Rightarrow n_i^2 = n_i$$

CLT says $(\frac{n}{S})$ is gaussian distributed as $S$ gets large:

$$\rho\left(\tfrac{n}{S}\right) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \exp\left[\frac{-\left(\frac{n}{S} - \left\langle \frac{n}{S} \right\rangle\right)^2}{2\tilde{\sigma}^2}\right]$$

where

$$\begin{aligned}
\tilde{\sigma}^2 &= \frac{1}{S^2}\sum_{i=1}^{S}\left(\left\langle n_i^2 \right\rangle - \langle n_i \rangle^2\right) \\
&= \frac{1}{S^2}\left(Sp - Sp^2\right) \\
&= \frac{p}{S}(1-p) \\
&:= \frac{\sigma}{S} \\
\Rightarrow \sigma^2 &= p(1-p)
\end{aligned}$$

Changing variables:

$$\rho\left(\frac{n}{S}\right)d\left(\frac{n}{S}\right) = \rho(n)dn$$

$$\begin{aligned}\Rightarrow \rho(n) = \frac{1}{S}\rho\left(\frac{n}{S}\right) &= \frac{1}{\tilde{\sigma}\sqrt{2\pi}S}\exp\left[\frac{-(n-\langle n\rangle)^2}{2S^2\tilde{\sigma}^2}\right] \\ &= \frac{1}{\sigma\sqrt{2\pi S}}\exp\left[\frac{-(n-\langle n\rangle)^2}{2S\sigma^2}\right]\end{aligned}$$

Let $\tilde{\tilde{\sigma}} = \sqrt{S}\sigma$:

$$= \frac{1}{\tilde{\tilde{\sigma}}\sqrt{2\pi}}\exp\left[\frac{-(n-\langle n\rangle)^2}{2\tilde{\tilde{\sigma}}^2}\right]$$

$$\tilde{\tilde{\sigma}}^2 = s\sigma^2 = Sp(1-p)$$

And we see that $n$ is indeed gaussian distributed.

## 1.3 Review of Thermodynamics

### 1.3.1 Statistical Description of Large Systems

**Problem**: From §1.1.5, we saw that for macroscopic systems ($\sim 10^{23}$ particles), we can not solve deterministically by solving the equations of motion, even if this were desirable.

**Solution**: Distinguish between a system's microstate, which is classically given by a point $\gamma \in \Gamma$ (Quantum mechanically, this is a stationary state given by a many body wavefunction $\Psi(x_1, x_2, \ldots, x_{10^{23}})$) and its macrostate, which is specified by values of certain macroscopic parameters (ie. energy, temperature, pressure, etc.).

**Definition 1:**

- For a given energy $E$, the number of microstates whose energy is $\leq E$ is denoted by $\Omega(E)$ and is called the integrated density of states or the total number of states.
- The density of states is

  $$\omega(E) := \frac{d\Omega(E)}{dE}$$

  which is the number of microstates per energy interval at a given point in energy space.

  $$\Rightarrow \Omega(E) = \omega(E)\,dE$$

  where $\Omega(E)$ depends on the size of $dE$, but $\omega(E)$ is independent of the increment size.

- For a system whose energy is constrained to the interval $[E - \Delta E, E]$, the number of microstates with energies in the interval are denoted by $\Omega_{\Delta E}(E)$, and is called the number of accessible states.

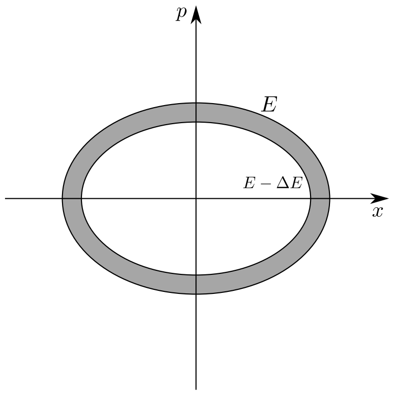
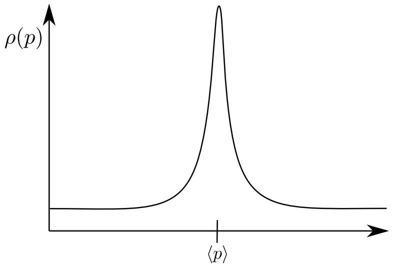


Figure 1.4: Example with Classic Harmonic Oscillator

***Remark 1:*** For macroscopic systems, any reasonable choice of $\Delta E$ has $\Omega(E) \approx \Omega_{\Delta E}(E)$ to tremendous accuracy. Thus it is often unnecessary to distinguish between $\Omega(E)$ and $\Omega_{\Delta E}(E)$.

***Remark 2:*** For large systems ($N$ particles, $N \gg 1$), $\Omega(E)$ increases exponentially with $N$, and for fixed $\frac{E}{N} = const$, then $\Omega(E)$ and $\omega(E)$ are rapidly growing functions of energy.

***Remark 3:*** The goal of a statistical description is to determine probability distributions for the values of an observable system that is characterized by certain macroscopically defined conditions.



### 1.3.2 The Equilibrium State

**Definition 1:** A system is isolated if it can not exchange energy with its surroundings (except for infinitesimal perturbations due to the butterfly effect).

***Remark 1:*** A completely isolated system would, in general, behave very differently from the realistic "isolated" system above.

**Definition 2:** A system is in equilibrium if the probability of finding the system in a particular microstate does not change with time.

***Remark 2:*** In equilibrium, all probability distributions are time independent, and thus all average values of observables are time independent as well. However, the actual values still can fluctuate (equilibrium is *not* a static state).

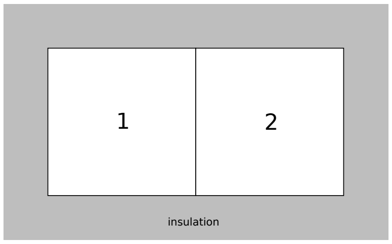
**Postulate of Equal a priori Probabilities**: An isolated system in equilibrium is equally likely to be found in any of its accessible states.

**Postulate - An Approach to Equilibrium**: An isolated system that is not in equilibrium will approach equilibrium if left undisturbed for a sufficiently long time.

***Remark 3:*** The above postulate would be false for truly isolated systems.

### 1.3.3 Interacting Systems

Two systems that are in contact with one another but isolated from the rest of the world can exchange energy in the form of heat ($Q$) and/or work ($W$). Thus the energy of each subsystem is not fixed.



**Definition 1:** The average energy of a system that is a subsystem of a larger isolated system is called its internal energy, and is denoted by $U \equiv \langle E \rangle$.

***Remark 1:*** If a system absorbs heat $Q$ and does work $W$, then its internal energy changes $\Delta U = Q - W$. (1st law of thermo)

***Remark 2:*** If the change is infinitesimal, then

$$dU = \delta Q - \delta W$$

where $dU$ is an exact differential while $\delta Q$ and $\delta W$ are *not*.

***Remark 3:*** $\Delta U$ depends only on the systems initial and final states, while $Q$ and $W$ also depend on the process that takes the system from the initial to final state.

### 1.3.4 Reversible vs Irreversible Processes

**Definition 1:**

- A quasi-static process is one that is infinitesimally slow enough so that the system remains in equilibrium at all times.
- An adiabatic process does not involve heat transfer.
- A reversible process is both quasi-static and adiabatic.

***Remark 1:*** Let $\Omega_i$ and $\Omega_f$ be the initial and final number of accessible states in a process. For processes that are adiabatic and quasi-static, $\Omega_i = \Omega_f$, and are thus reversible.

***Remark 2:*** Let $x$ be an external parameter and let $E_n(x)$ be the energy of the n$^{th}$ microstate as a function of $x$. If $x$ changes quasi-statically, then the sytem does work:

$$\delta W = -\left\langle \frac{\partial E_n}{\partial x} \right\rangle dx := F\, dx$$

where

$$F := -\left\langle \frac{\partial E_n}{\partial x} \right\rangle$$

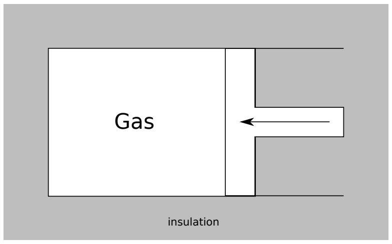
and is called the generalized force conjugate to $x$.

**Example**: If the system volume changes by an amount $dV$, then work is done:

$$\delta W = p\, dV, \qquad \text{with} \quad p = -\left\langle \frac{dE}{dV} \right\rangle$$

where $p =$ pressure is the generalized force conjugate to the volume.

**Example**: The infinitely slow compression of a gas in a thermally insulated cylinder by a thermally insulated piston is reversible.



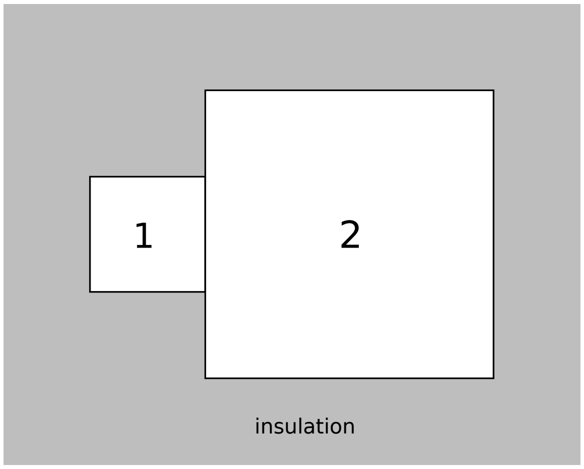
**Definition 2:** An irreversible process is one that increases the number of accessible states such that $\Omega_f > \Omega_i$.

***Remark 3:*** $\Omega_f > \Omega_i$ implies that the final state has a higher probability and thus the system will not spontaneously return to the initial state.

**Example**: Humpty Dumpty or the free expansion of a gas into a vacuum are examples of irreversible processes.

### 1.3.5 Energy, Temperature, and Entropy

Consider a "small" system (1), and a "large" system (2) that are in thermal contact. Let their energies be $E_1$ and $E_2$ and the total energy $E_{tot} = E_1 + E_2$ and their number of accessible states are $\Omega_1$ and $\Omega_2$. Let $E_1 = E$, then the total number of accessible states for the whole system:



$$\Omega_{tot}(E) = \Omega_1(E)\Omega_2(E_{tot} - E)$$

***Remark 1:*** The total number of states for the whole system is

$$\Omega_{tot} = \sum_E \Omega_{tot}(E)$$

and the probability distribution for $E$ is

$$\rho(E) = \frac{\Omega_{tot}(E)}{\Omega_{tot}}$$

***Remark 2:*** Recall from §1.3.1.R2, we expect $\Omega_{tot}(E)$ to be sharply peaked about $\langle E\rangle = U$, the internal energy of system 1.

**Definition 1:** A system's entropy is given by:

$$S := k_B \ln\left(\Omega(E)\right)$$

where $k_B = 1.38 \times 10^{-23}\frac{J}{K}$ = Boltzmann's constant.

***Remark 3:*** For two combined systems, $\Omega = \Omega_1\Omega_2 \;\Rightarrow\; S = S_1 + S_2$, and thun entropy is additive.

***Remark 4:*** The entropy change for any process that takes an isolated system from one macrostate to another is positive semi-definite: $\Delta S \geq 0$, where the equality only holds if the process is reversible.

**Definition 2:** The temperature, $T$, of a system is defined as

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{E=U} = k_B \left.\frac{\partial(\ln\Omega)}{\partial E}\right|_{E=U}$$

Also, let

$$\beta := \frac{1}{k_B T}$$

***Remark 5:*** For two systems in equilibrium, the equilibrium state (most probable) occurs when

$$\Omega_1(E)\Omega(E_{tot} - E)$$

is maximized. Which also implies

$$S_1(E) + S_2(E_{tot} - E)$$

is maximized. Taking the derivatives to find the max:

$$\frac{\partial}{\partial E}\left[S_1(E) + S_2(E_{tot} - E)\right] = \frac{\partial S_1}{\partial E} - \frac{\partial S_2}{\partial E} = 0$$

$$\Rightarrow T_1 = T_2 \qquad \text{in equilibrium}$$

***Remark 6:*** If $x$ is the value of some external parameter, then let $\Omega(E, x)$ be the corresponding number af accessible states. Then the generalized force conjugate to $x$ is given by

$$F = \frac{1}{\beta}\frac{\partial \ln \Omega}{\partial x}$$

***Remark 7:*** The relation between temperature, any external parameter, and the generalized conjugate forces,

$$T = T(x, F)$$

is called the equation of state.

***Remark 8:*** Knowledge of $\Omega(E, x)$ or, equivalently, $S(E, x) = k_B \ln \Omega(E, x)$, is sufficient for calculating all the thermodynamic properties. In particular:

$$\beta = \left.\frac{\partial \ln \Omega}{\partial E}\right|_{E=U} ; \qquad \frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{E=U}$$

$$F = \frac{1}{\beta}\left.\frac{\partial \ln \Omega}{\partial x}\right|_{x=\langle x\rangle} = T\left.\frac{\partial S}{\partial x}\right|_{x=\langle x\rangle}$$

### 1.3.6 The Laws of Thermodynamics

**First Law**:

- For an isolated system, $U$ is a constant.
- If a system is brought from one macrostate to another by a process that causes the system to absorb heat $Q$ and do work $W$, then the internal energy changes by $\Delta U = Q - W$.

***Remark 1:*** The first law simply expresses energy conservation.

**Second Law**:

- For any process that can take an isolated system from one macrostate to another, $\Delta S \geq 0$.
- A quasi-static process during which a system absorbs an infinitesimal amount of heat $\delta Q$ gives an entropy change $dS = \frac{\delta Q}{T}$.

***Remark 2:*** $dS$ is an exact differential, but $\delta Q$ is not.

***Remark 3:*** The 2nd law has profound technical consequences, which will be covered later in the semester.

***Remark 4:*** Combining the first and second laws gives

$$dU = T\,dS - \delta W$$

If the work done is purely mechanical, then $\delta W = p\,dV$

$$\Rightarrow dU = T\,dS - p\,dV$$

**Third Law**: $S(T \to 0) \to 0$ (Also called Nerst's Law)

***Remark 5:*** Dropping the distinction between $\Omega(E)$ and $\Omega_{\Delta E}(E)$ is *not* compatible with the 3rd law. More generally, entropy can not be defined in a classical context (more on that later!).

### 1.3.7 Thermodynamic Potentials

**Definition 1:** In addition to internal energy, we define:

- Enthalpy: $H = U + pV$
- Helmholtz Free Energy: $F = U - TS$
- Gibb's Free Energy: $G = H - TS$

***Remark 1:*** These are obtained from Legendre transforms of the 1st law. For example:

$$\begin{aligned} dU &= T\,dS - p\,dV \\ &= T\,dS - d(pV) + V\,dp \\ d(U + pV) &= T\,dS + V\,dp \\ &= dH \end{aligned}$$

Similar transforms give $F$ and $G$. These give the differentials:

$$\begin{aligned} dH &= T\,dS + V\,dp \\ dF &= S\,dT - p\,dV \\ dG &= -S\,dT + V\,dp \end{aligned}$$

***Remark 2:*** These relations imply

$$T = \left(\frac{\partial H}{\partial S}\right)_p = \left(\frac{\partial U}{\partial S}\right)_V$$
$$p = -\left(\frac{\partial F}{\partial V}\right)_T = -\left(\frac{\partial U}{\partial V}\right)_S$$

etc. They also imply Maxwell's Relations:

$$\Rightarrow \left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T$$
$$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p$$

***Remark 3:*** The specific heats are given by:

$$C_v = \left(\frac{\delta Q}{dT}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V$$
$$C_p = \left(\frac{\delta Q}{dT}\right)_p = T\left(\frac{\partial S}{\partial T}\right)_p = \left(\frac{\partial H}{\partial T}\right)_p$$

### 1.3.8 The Simple Example: The Ideal Gas

A monatomic ideal gas with $N$ particles has equation of state:

$$PV = nk_BT = nRT$$

It has internal energy

$$U = \frac{3}{2}Nk_BT$$

and specific heat / mol:

$$C_V = \frac{3}{2}R$$
$$C_p = \frac{5}{2}R$$

Note that these are only valid for monatomic gas. Now, a quasi-static change from $(T_i, V_i)$ to $(T_f, V_f)$ gives an entropy change of:

$$\Delta S = nC_V \ln\left(\frac{T_f}{T_i}\right) + nR\ln\left(\frac{V_f}{V_i}\right)$$

***Remark 1:*** All of this can be derived from

$$\Omega(E) = V^N \chi(E)$$

where $\chi(E)$ is not an explicit function of $V$.

## 1.4 Statistical Ensembles

### 1.4.1 Gibbsian Ensembles

Consider a system in a particular microstate.

**Definition 1:** A statistical ensemble consists of many identical systems (same number of particles governed by the same equations of motion) that are all in the same macrostate but, in general, in different microstates.

***Remark 1:*** This concept makes averaging easier since we can average over members of the ensemble at a given time rother than following the time evolution of a syngle system and averaging over long time periods.

**Example**: Probability of getting a 6 when rolling 1 die once. You can either roll one die many times, or roll many identical dice at once.

***Remark 2:*** Need to find the probability of a member of the ensemble to be in a certain microstate. This depends on the physical situation.

### 1.4.2 The Microcanonical Ensemble

Consider an isolated system in equilibrium whose microstates have energies

$$E(p,q) \qquad \text{for the classical case, or}$$

$$E_n, \quad n = 0, 1, \ldots \qquad \text{for the quantum mechanical case}$$

Define a macrostate by demanding thath the energy lies in an interval $[E - \Delta E, E]$.

**Theorem**: The probability of the system being in the microstate with energy $E(p,q)$ or $E_n$ is:

$$\rho(p,q) = \frac{1}{\int d\gamma} \cdot \begin{cases} 1 & \text{if } E - \Delta E \le E(p,q) \le E \\ 0 & \text{otherwise} \end{cases} \tag{$\star$}$$

where the integral is taken over the region where $E - \Delta E \le E(p,q) \le E$ and

$$\gamma \equiv (p,q), \qquad d\gamma = dp\,dq, \qquad \Gamma = \{\gamma\}$$

$$\rho(E_n) = \frac{1}{\sum 1} \cdot \begin{cases} 1 & \text{if } E - \Delta E \le E_n \le E \\ 0 & \text{otherwise} \end{cases} \tag{$\star$}$$

Proof: Follows from §1.3.2.P1.

***Remark 1:***

$$\int \rho(p,q)\, d\gamma = 1; \qquad \sum_n \rho(E_n) = 1$$

And so our normalization is ok.

***Remark 2:*** This ensemble is called microcanonical and ($\star$) is called the microcanonical distribution function. It is appropriate for isolated systems.

***Remark 3:*** Microcanonical ensembles lack an obvious temperature concept.

### 1.4.3 The Canonical Ensemble

**Definition 1:** A heat bath is a system that is large enough so that any heat that is absorbed or given off by the heat bath in any thermodynamic process is negligible compared to its internal energy.

**Definition 2:** The canonical partition function, $Z$, is defined quantum mechanically as:

$$Z = \sum_n e^{-\beta E_n}$$

and classically, for a system with $N$ particles and $3N$ degrees of freedom, as:

$$Z = \frac{1}{\hbar^{2N} N!} \int e^{-\beta E(\gamma)}\, d\gamma$$

***Remark 1:*** The uncertainty relation implies that the phase space can not be partitioned into cells smaller than $\Delta p \Delta q = \hbar$. Thus the factor of $\hbar^{2N}$ in the classical definition. We'll discuss the origin of the $N!$ later.

***Remark 2:*** A more general definition of the partition function which holds classically and quantum mechanically is:

$$Z = \int_0^\infty \omega(E) e^{-\beta E}\, dE$$

where $\omega(E)$ is the density of states.

**Theorem**: For a system that is in thermal contact with a heat bath, the probability of being in a microstate with energy $E_n$ (QM) is:

$$\rho_n = \frac{1}{Z} e^{-\beta E_n} \qquad (\star\star)$$

or classically:

$$\rho_E = \frac{1}{\hbar^{3N} N! Z} e^{-\beta E} \qquad (\star\star)$$

where $\beta = \frac{1}{k_B T}$ and $T$ is the temperature of the heat bath.

Proof (QM only): Let 1 = system and let 2 = heat bath. Thus we have that 1 is in a state with energy $E_n$, and 2 has energy $E_2$. Thus, the combined system has

$$\begin{aligned} E_{tot} &= E_n + E_2 \\ &= \text{constant by 1st law} \end{aligned}$$

By §1.3.2.P1, we have that

$$\begin{aligned} \rho_n &= (\text{const})\Omega_2(E_{tot} - E_n) \\ \Rightarrow \rho_n &= (\text{const})\exp\left[\frac{S_2(E_{tot} - E_n)}{k_B}\right] \end{aligned}$$

Because of the heat bath, $E_{tot} \gg E_n$, so:

$$\begin{aligned} \rho_n &= (\text{const}')\exp\left[-E_n \underbrace{\left.\frac{\partial S_2}{\partial E_n}\right|_{E_{tot}} \frac{1}{k_B}}_{\beta} + \cdots\right] \\ &= (\text{const}')e^{-\beta E_n} \end{aligned}$$

Normalization implies

$$(\text{const}') = \left(\sum_n e^{-\beta E_n}\right)^{-1} = \frac{1}{Z}$$

***Remark 3:*** This is called a canonical ensemble and ($\star\star$) is a canonical distribution or a Gibbs-Boltzmann distribution.

***Remark 4:*** The canonical distribution also applies to situations where the system has been prepared so that its energy is not known precisely but the average energy, $U$, is known. In this case, $T$ is the temperature a heat bath needs to have in order for the average energy to remain $U$ if the system is brought into contact with the heat bath.

***Remark 5:*** The energy of a system in contact with a heat bath is *not* fixed, but is subject to fluctuations. How big are these fluctuations?

**Lemma**: In a canonical ensemble, the mean energy and the rms energy fluctuations are determined by the partition function via the following relationships:

$$U = \langle E\rangle = -\frac{\partial \ln Z}{\partial \beta}$$

$$(\Delta E) = \left\langle (E - \langle E \rangle)^2 \right\rangle^{1/2} = \left( \frac{\partial^2 \ln Z}{\partial \beta^2} \right)^{1/2}$$

Proof:

$$Z = \sum_n e^{-\beta E_n}$$

Thus

$$\begin{aligned} -\frac{\partial \ln Z}{\partial \beta} &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \\ &= \frac{1}{Z} \sum_n E_n e^{-\beta E_n} \\ &= \sum_n E_n \rho_n \\ &= \langle E \rangle \end{aligned}$$

And

$$\begin{aligned} \frac{\partial^2 \ln Z}{\partial \beta^2} &= \frac{\partial}{\partial \beta} \left[ -\frac{1}{Z} \sum_n E_n e^{-\beta E_n} \right] \\ &= \frac{1}{Z^2} \left( \frac{\partial Z}{\partial \beta} \right) \sum_n E_n e^{-\beta E_n} + \frac{1}{Z} \sum_n E_n^2 e^{\beta E_n} \\ &= - \langle E \rangle \frac{1}{Z} \sum_n E_n e^{-\beta E_n} + \left\langle E^2 \right\rangle \\ &= - \langle E \rangle^2 + \left\langle E^2 \right\rangle \\ &= (\Delta E) \end{aligned}$$

***Remark 6:*** The heat capacity, $C_V = \left(\frac{dU}{dT}\right)_V$:

$$\Rightarrow C_V = -\frac{\beta}{T} \left( \frac{\partial U}{\partial \beta} \right)_V = \frac{\beta}{T} (\Delta E)^2$$

***Remark 7:*** For a macroscopic system with $N$ particles at temperature $T$,

$$U = \mathcal{O}\left(N k_B T\right)$$

$$C_V = \mathcal{O}\left(N k_B\right)$$

**Proposition 1**: The energy fluctuations in a large system described by a canonical ensemble are gaussian, and the relative energy fluctuations are:

$$\frac{\Delta E}{\langle E \rangle} = \mathcal{O}\left( \frac{1}{\sqrt{N}} \right)$$

with $N$ equalling the number of particles. So the fluctuations are small, and thus we need only worry about the averages.
Proof: Consider a large system, and divide it up into $n$ identical subsystems that are each still macroscopic (and $n \gg 1$). Let each subsystem have energy $E_i$, $(i = 1, \ldots, n)$. Thus, the total energy of the system is:

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | ⋯ | |
| | | | | |
| | | | | |
| | | | | |
| | | | ⋯ | n |

$$E_{tot} = \sum_{i=1}^{n} E_i$$

Due to fluctuations, $E_i$ are random variables that obey the canonical distribution. Thus each subsystem has a canonical partition function:

$$z_i = \sum_m e^{-\beta E_m^{(i)}}$$

where $E_m^{(i)}$ is the energy of subsystem number $i$. The subsystems are identical, and thus $z_i = z$. Thus, the partition function of the whole system can be written:

$$Z = z^n$$

The canonical distribution has a finite variance, and hence CLT applies. By §1.2.6, we have:

$$x_i = nE_i \quad (i = 1, \ldots, n)$$

$$y = \frac{1}{n} \sum_{i=1}^{n} x_i = E$$

$$\Rightarrow \rho(E) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(E - \langle E \rangle)^2}{2\sigma^2}\right]$$

with

$$\sigma^2 = \sum_{i=1}^{n} \left( \left\langle E_i^2 \right\rangle - \langle E_i \rangle^2 \right) = n \left\langle (E - \langle E \rangle)^2 \right\rangle = n(\Delta E)^2$$

By the above Lemma:

$$\sigma^2 = \frac{n\partial^2 \ln z}{\partial \beta^2} = \frac{\partial^2 \ln Z}{\partial \beta^2} = (\Delta E_{tot})^2 = \frac{T}{\beta} C_V$$

$$\Rightarrow \frac{(\Delta E_{tot})^2}{\langle E\rangle^2} = \frac{\sigma^2}{U^2} = \frac{1}{U^2}\frac{T}{\beta}C_V = \mathcal{O}\left[\frac{T}{Nk_BT}\frac{k_BT}{Nk_BT}Nk_B\right] = \mathcal{O}\left[\frac{1}{N}\right]$$

***Remark 8:*** A macroscopic system has an extremely sharp gaussian energy distribution.

**Proposition 2**: In a canonical ensemble, the Helmholtz free energy is given by:

$$F = -k_BT\ln Z$$

<u>Proof:</u> Let $x$ be an external parameter that characterizes the system. Then the energy level $E_n = E_n(x)$.

$$\Rightarrow Z = \sum_n e^{-\beta E_n(x)} = Z(\beta, x)$$

If we change $x, \beta$ quasi-statically by $dx$ and $d\beta$ respectively, then

$$\begin{aligned} d\ln Z &= \frac{\partial \ln Z}{\partial \beta}d\beta + \frac{\partial \ln Z}{\partial x}\,dx \\ &= -U\,d\beta + \frac{1}{Z}\sum(-\beta)e^{-\beta E_n(x)}\frac{\partial E}{\partial x}\,dx \\ &= -U\,d\beta - \beta\left\langle\frac{\partial E_n}{\partial x}\right\rangle\,dx \end{aligned}$$

By §1.3.4.R2:

$$\begin{aligned} &= -U\,d\beta + \beta\,\delta W \\ &= \beta\,\delta W - d(U\beta) + \beta\,dU \\ &= \beta\,(\delta W + dU) - d(U\beta) \\ &= \beta\,(\delta Q) - d(U\beta) \\ &= \frac{1}{k_BT}T\,dS - d(U\beta) \\ \Rightarrow d(\ln Z + U\beta) &= \frac{1}{k_B}\,dS \end{aligned}$$

$$\Rightarrow S = k_B\,(\ln Z + \beta U) = k_B\ln Z + \frac{U}{T} \qquad \text{(up to a constant)}$$

Thus by the definition of $F$:

$$\begin{aligned} F &= U - TS \\ &= -k_BT\ln Z \end{aligned}$$

***Remark 9:*** $F$, and therefore $Z$, determine all thermodynamic quantities.

**Corollary 1**: Up to a constant,

$$S = k_B \ln Z + k_B T \frac{\partial \ln Z}{\partial T}$$

Proof: By §1.3.7:

$$S = -\frac{\partial F}{\partial T} = k_B \ln Z + k_B T \frac{\partial \ln Z}{\partial T}$$

**Corollary 2:**:

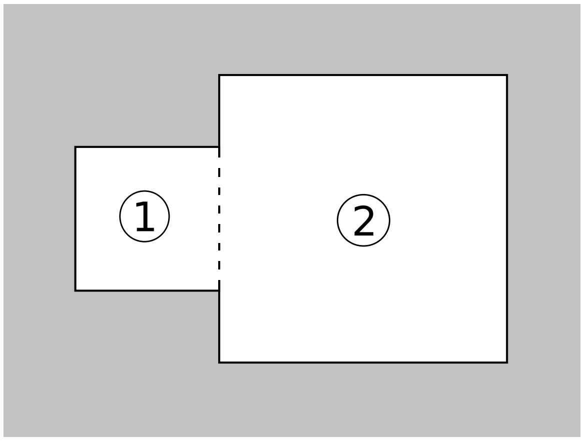
$$p = k_B T \frac{\partial \ln Z}{\partial V}$$

Proof: Also by §1.3.7:

$$p = -\frac{\partial F}{\partial V} = \frac{\partial}{\partial V}\left(k_B T \ln Z\right) = k_B T \frac{\partial \ln Z}{\partial V}$$

### 1.4.4 The Grand Canonical Ensemble

Consider System ① in contact with a heat bath (System ②):



In this case, due to the perforated boundry, both heat and particles can be exchanged.

***Remark 1:*** The particle number $N$ is not fixed, and so the number of accessible states depends on both $N$ and $F$.

$$\Rightarrow \Omega = \Omega(E, N)$$

**Definition 1:** The chemical potential, $\mu$, of a system is

$$\mu := -\frac{1}{\beta} \left.\frac{\partial \ln \Omega}{\partial N}\right|_{\substack{E=\langle E\rangle=U \\ N=\langle N\rangle}}$$

***Remark 2:*** $-\beta\mu$ is related to $N$ as $\beta$ is related to energy (§1.3.5.R8):

$$\beta = \left.\frac{\partial \ln \Omega}{\partial E}\right|_{E=U}, \qquad -\beta\mu = \left.\frac{\partial \ln \Omega}{\partial N}\right|_{\substack{E=U \\ N=\langle N\rangle}}$$

***Remark 3:*** Many properties of $\beta$ carry over to $\mu$. In particular, two systems in equilibrium with one another have the same $\mu$. In addition, a heat bath will have constant $\mu$.

**Definition 2:** Let $E_n^{(N)}$ $(n = 0, 1, \ldots)$ be the energy levels of ① when it contains $N$ particles. The grand canonical partition function is defined as

$$Z = \sum_{N=0}^{\infty} e^{\beta\mu N} \sum_{n=0}^{\infty} e^{-\beta E_n^{(N)}}$$

***Remark 4:*** Here we restrict ourselves to quantum mechanics.

**Theorem**: The probability of system ① being in a macrostate with $N$ particles and energy $E_n^{(N)}$ is

$$\rho_n^{(N)} = \frac{1}{Z}\exp\left[-\beta\left(E_n^{(N)} - \mu N\right)\right]$$

Proof: Generalizing §1.4.3, we have that

$$E_n^{(N)} + E_2 = E_{tot}$$

$$N + N_2 = N_{tot}$$

$$\begin{aligned} S_2(E_2, N_2) &= k_B \ln \Omega_2 \left(E_{tot} - E_n^{(N)}, N_{tot} - N\right) \\ &\approx k_B \ln \Omega_2 \left(E_{tot}, N_{tot}\right) - k_B E_n^{(N)} \left.\frac{\partial \ln \Omega_2}{\partial E}\right|_{\substack{E_{tot} \\ N_{tot}}} - k_b N \left.\frac{\partial \ln \Omega_2}{\partial N}\right|_{\substack{E_{tot} \\ N_{tot}}} \\ &= \mathcal{C} - k_B \beta E_n^{(N)} + k_B N \beta\mu \end{aligned}$$

$$\Rightarrow \rho_n^{(N)} \sim \Omega_2 = e^{S_2/k_B} = \mathcal{C}' e^{-\beta E_n^{(N)}} e^{\beta\mu N}$$

where $\mathcal{C}$ in an constant. Normalizing, we find that

$$\mathcal{C}' = \frac{1}{Z}$$

**Lemma**: In a grand canonical ensemble, the thermodynamic potentials $U, H, F, G$ obey the relations:

$$\begin{aligned} dU &= T\,dS - p\,dV + \mu\,dN, && U(S,V,N) \\ dH &= T\,dS + V\,dp + \mu\,dN, && H(S,p,N) \\ dF &= -S\,dT - p\,dV + \mu\,dN, && F(T,V,N) \\ dG &= -S\,dT + V\,dp + \mu\,dN, && G(T,p,N) \end{aligned}$$

<u>Proof:</u> Recall that

$$\mu = -k_B T \left.\frac{\partial \ln \Omega}{\partial N}\right|_{\substack{E=U \\ N=\langle N\rangle}}$$

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

But then

$$\begin{aligned} dS &= \left.\left(\frac{\partial S}{\partial E}\right)_{V,N}\right|_{E=U} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \\ &= \frac{1}{T}\,dU + \frac{p}{T}\,dV - \frac{\mu}{T}\,dN \end{aligned}$$

$$\Rightarrow dU = T\,dS - p\,dV + \mu\,dN$$

Performing Legendre transformations, it follows that:

$$\begin{aligned} H = U + PV &\Rightarrow dH = T\,dS + V\,dp + \mu\,dN \\ F = U - TS &\Rightarrow dF = -S\,dT - p\,dV + \mu\,dN \\ G = F + pV &\Rightarrow dG = -S\,dT + V\,dp + \mu\,dN \end{aligned}$$

***Remark 5:*** $\mu$ is the energy change that results from adding one particle to the system adiabatically ($dS = 0$) and isochorically ($dV = 0$).

**Lemma 2**: The Duham-Gibbs Relation:

$$G = N\mu$$

<u>Proof:</u> By Lemma 1:

$$G = G(T,p,N)$$

We expect $G$ to be proportional to $N$, and so we let

$$G(T,p,N) = Ng(T,p) \tag{$\star$}$$

Now, again by Lemma 1:

$$\mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} = g(T,p) = \frac{G}{N}$$

$$\Rightarrow G = \mu N$$

***Remark 6:*** It is not obvious that ($\star$) holds. See §1.4.5 for a discussion and derivation of ($\star$).

**Proposition 1**: In a grand canonical ensemble:

$$J := -k_B T \ln Z$$

defines a thermodynamic potential whose differential is

$$dJ = -S\,dT - p\,dV - N\,d\mu$$

which is also equal to

$$J = -pV$$

This is called the grand potential.

Proof: Perform a Legendre transformation from $F = F(T,V,N)$ to $J = F - \mu N = J(T,V,\mu)$.

$$\begin{aligned}\Rightarrow dJ &= dF - \mu\,dN - N\,d\mu \\ &= -S\,dT - p\,dV - N\,d\mu\end{aligned}$$

From Lemma 2 ($G = \mu N$):

$$\Rightarrow J = F - \mu N = F - G = F - (F + pV) = -pV$$

We must still show that:

$$J = -k_B T \ln Z$$

Now

$$\begin{aligned}\rho_n^{(N)} &= \frac{1}{Z}\exp\left[-\beta E_n^{(N)} + \beta\mu N\right] \\ &\sim \Omega_2\left(E_{tot} - E_n^{(N)}, N_{tot} - N\right) \\ &\sim \frac{\Omega_{tot}}{\Omega_1(E_n^{(N)}, N}\end{aligned}$$

$$\Rightarrow \ln \rho_n^{(N)} = \mathcal{C} + \ln \Omega_{tot} - \ln \Omega_1\left(E_n^{(N)}, N\right)$$

$$\begin{aligned}\Rightarrow S &= k_B \left\langle \ln \Omega_1 \left( E_n^{(N)}, N \right) \right\rangle = - \left\langle \ln \rho_n^{(N)} \right\rangle \quad \text{up to a const} \\ &= k_b \ln Z + k_B \beta \left\langle E_n^{(N)} \right\rangle - k_B \beta \mu \left\langle N \right\rangle\end{aligned}$$

Multiplying everything by $(-T)$ yields

$$\begin{aligned}-k_B T \ln Z &= TS + U - \mu \left\langle N \right\rangle \\ &= F - \mu \left\langle N \right\rangle \\ &= J\end{aligned}$$

**Proposition 2**: In a grand canonical ensemble, the mean particle number is given by

$$\left\langle N \right\rangle = k_B T \frac{\partial \ln Z}{\partial \mu}$$

$$Z = \sum_N e^{\beta \mu N} \sum_n e^{-\beta E_n^{(N)}}$$

Proof: See HW.

### 1.4.5 The Thermodynamic Limit and the Durham-Gibbs Relation

Consider a macroscopic system, $V = \mathcal{O}(1 \text{ cm}^3, N = \mathcal{O}(10^{20})$, which is huge compared to its microscopic dimensions of $N \sim \mathcal{O}(1), V \sim \mathcal{O}\left(1\text{Å}^3\right)$. Recall that the free energy is $F = F(T, V, N)$.

**Definition 1:** The limit as $V \to \infty$ and $N \to \infty$ so that $n = \frac{N}{V} = \mathcal{C}$ and $f = \frac{F}{V} = \mathcal{C} < \infty$ is called the thermodynamic limit.

***Remark 1:*** For practical purposes, macroscopic systems are in the thermodynamic limit.

**Definition 2:**

- Consider a function of $n$ variables, $f(x_1, \ldots, x_n)$. It is a homogeneous function of degree $a$ if the number $a \in \mathbb{R}$ exists so that

$$f(bx_1, \ldots, bx_n) = b^a f(x_1, \ldots, x_n) \quad \forall b > 0$$

- $f$ is a generalized homogeneous function if $n+1$ numbers $[x_1], [x_2], \ldots, [x_n], [f]$ (not all zero) exist so that for all $b > 0$

$$f(b^{[x_1]}x_1, b^{[x_2]}x_2, \ldots, b^{[x_n]}x_n) = b^{[f]} f(x_1, x_2, \ldots, x_n)$$

**Examples**:

1. If we have the case where $n = 1, \quad f(x) = cx^a$

   $$\Rightarrow f(bx) = c(bx)^a = b^a cx^a = b^a f(x)$$

   $$\Rightarrow \text{ homogeneous function of degree } a$$

   If instead we had $n = 1, \quad f(x) = \ln(x)$:

   $$\Rightarrow f(bx) = \ln(bx) = \ln b \ln x$$

   $$\Rightarrow \text{ Not homogeneous}$$

2. Now consider when $n = 2, \quad f(x, y) = x^3 + xy + y^{3/2}$. Let's choose $[x] = 1$, $[y] = 2$.

   $$\begin{aligned}\Rightarrow f(b^{[x]}x, b^{[y]}y) &= f(bx, b^2y)\\ &= (bx)^3 + (bx)(b^2y) = (b^2y)^{3/2}\\ &= b^3\left(x^3 + xy + y^{3/2}\right)\\ &= b^3 f(x, y)\end{aligned}$$

   $$\Rightarrow \quad f(x, y) \text{ is a generalized homogeneous function with } [f] = 3.$$

   What is we hadn't known what $[x]$ and $[y]$ were?

   $$x^3 b^{3[x]} + b^{[x]+[y]}xy + b^{3[y]/2}y^{3/2} = b^{[f]}(x^3 + xy + y^{3/2})$$

   $$\Rightarrow [x] + [y] = 3[x] = \frac{3}{2}[y] = [f]$$

   So we basically have 3 equations, 3 unknowns. The solution will be unique up to a redefinition of $b$. As a counter example, consider:

   $$n = 2: \qquad f(x, y) = x^3 + e^y \text{ is } \textit{not} \text{ a generalized homogeneous function}$$

Consider now a $d$-dimensional system in the thermodynamic limit, and change all linear dimensions by a factor of $b > 0$.

$$L \to L' = \frac{L}{b}$$

$$V \to V' = b^{-d}V$$

***Remark 2:*** Under this transformation, in order for $n = \frac{N}{V}$ to remain constant, we must have

$$N \to N' = b^{-d}N$$

***Remark 3:*** If the system is in contact with a heat bath, where the system is huge in comparison to the microscopic limit, and the heat bath is huge in comparison to the system, then $T$ does not change upon changing volume.

$$T \to T' = b^0 T$$

**Lemma**: The Helmholtz free energy is a generalized homogeneous function of $T, V, N$, namely:

$$F(T,V,N) = b^d F\left(T, V b^{-d}, N b^{-d}\right)$$

Proof: By Definition 1:

$$\begin{aligned}\frac{F(T',V',N'}{V'} &= f(T',V',N') \\ &= f\left(T,V,N\right) \quad \text{in the thermodynamic limit} \\ &= \frac{F\left(T,V,N\right)}{V}\end{aligned}$$

$$\begin{aligned}\Rightarrow F\left(T,V,N\right) &= \frac{V}{V'} F\left(T',V',N'\right) \\ &= b^d F\left(T, b^{-d}V, b^{-d}N\right)\end{aligned}$$

***Remark 4:*** $[f] = -d$, $[V] = -d$, $[N] = -d$, $[T] = 0$ are called scale dimensions of $F, V, N, T$ respectively.

***Remark 5:*** Let $A$ and $B$ have scale dimensions of $[A]$ and $[B]$. Then $[AB] = [A] + [B]$.

**Lemma 2**: There exists a function $\tilde{F}(x,y)$ such that

$$F(T,V,N) = N\tilde{F}\left(T, \frac{V}{N}\right)$$

Proof: The value of $b$ in Lemma 1 is arbitrary so long as it remains positive. Thus, lets let $b = N^{1/d}$.

$$\Rightarrow F(T,V,N) = NF(T, N^{-1}V, 1)$$

Define $F\left(x,y,1\right) := \tilde{F}(x,y)$.

$$\Rightarrow F\left(T,V,N\right) = N\tilde{F}\left(T, \frac{V}{N}\right)$$

**Lemma 3**:

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

Proof: By Lemma 2:

$$\begin{aligned}\left(\frac{\partial F}{\partial N}\right)_{V,T}&=\tilde{F}\left(T,\frac{V}{N}\right)+N\left(\frac{-V}{N^2}\right)\left.\frac{\partial\tilde{F}(x,y)}{\partial y}\right|_{\substack{x=T\\y=V/N}}\\&=\tilde{F}\left(T,\frac{V}{N}\right)-\frac{V}{N}\left(\frac{\partial\tilde{F}(x=T,y=V/N)}{\partial y}\right)\end{aligned}$$

Since

$$\left(\frac{\partial F}{\partial N}\right)_{V,T}=N\frac{1}{N}\frac{\partial\tilde{F}}{\partial y}$$

$$\Rightarrow N\left(\frac{\partial F}{\partial N}\right)_{V,T}+V\left(\frac{\partial F}{\partial V}\right)_{N,T}=N\tilde{F}\left(T,\frac{V}{N}\right)=F\left(T,V,N\right)$$

**Theorem**: Again, $G=\mu N$ by the Duham-Gibbs Relation.

Proof: By §4.4.L1 $\Rightarrow dF=-S\,dT-p\,dV+\mu\,dN$.

$$\Rightarrow p=-\left(\frac{\partial F}{\partial V}\right)_{T,N},\qquad \mu=\left(\frac{\partial F}{\partial N}\right)_{T,V}$$

Now

$$\begin{aligned}G&=F+pV\\&=F-\left(\frac{\partial F}{\partial V}\right)_{T,N}V\\&=N\left(\frac{\partial F}{\partial N}\right)_{T,V}+V\left(\frac{\partial F}{\partial V}\right)_{T,N}-V\left(\frac{\partial F}{\partial V}\right)_{T,N}\qquad\text{by above L3}\\&=N\left(\frac{\partial F}{\partial N}\right)_{T,V}\\&=\mu N\end{aligned}$$

***Remark 6:*** The above proof did *not* use §4.4($\star$). ($G(T,p,N)=Ng(T,p)$)

***Remark 7:*** §4.4($\star$) can be obtained as follows:

$$[F]=[V]=-d$$

$$\Rightarrow [p] = \left[\frac{\partial F}{\partial V}\right] = \left[\frac{F}{V}\right] = [F] - [V] = 0$$

$$\Rightarrow G(T,p,N) = -b^{-[G]} G\left(Tb^{[T]}, pb^{[p]}, Nb^{[N]}\right)$$
$$= b^d G(T,p,Nb^{-d})$$

Let $b = N^{1/d}$:

$$\Rightarrow G(T,p,N) = NG(T,p,1) := Ng(T,p)$$

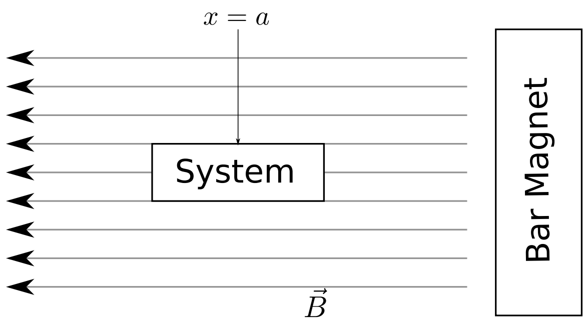
where $g(x,y) := G(x,y,1)$. Thus the proof of the Duham-Gibbs relation in §4.4 was indeed true, just not obvious.

### 1.4.6 Systems in Magnetic Fields

In general, work is not purely mechanical for systems in magnetic fields.

**Lemma**: A magnetization change $d\boldsymbol{M}$ by a magnetic field requires work $\delta W = \boldsymbol{B} \cdot d\boldsymbol{M}$.

Proof: We want to magnetize the system by letting it approach the bar magnet from $x = -\infty$. Let the field be axi-symmetric along the x-axis and $B(x \to \infty) \to 0$.



Also let the final position by $x = a$. The attractive force at point $x$ is:

$$F = M(x)\frac{dB}{dx}$$

To do this process quasi-statically, we need a balancing force $F_{balance} = -F$. So we want to apply force:

$$F_{balance}(x) = -M(x)\frac{dB}{dx}$$

which does work

$$\begin{aligned} W_1 &= -\int_{-\infty}^{a} M(x)\frac{dB}{dx}\,dx \\ &= -M(a)B(a) + \int_{-\infty}^{a} \frac{dM}{dx}B(x)\,dx \quad \text{(int by parts)} \\ &= -M(a)B(a) + \int_{0}^{M(a)} B(x)\,dM \end{aligned}$$

To find the work it takes to magnetize the system in the absence of any other changes, we now take the system back to $x = -\infty$ while keeping magnetization fixed at a value $M(a)$. This requires work.

$$\begin{aligned} W_2 &= -\int_{a}^{-\infty} M(a)\frac{dB}{dx}\,dx \\ &= -M(a)\,[0 - B(a)] \\ &= M(a)B(a) \end{aligned}$$

So the total required work to magnetize the system is:

$$W = W_1 + W_2 = \int_{0}^{M(a)} B(M)\,dM$$

For an infinitesimal magnetization then, the amount of work done on the system is

$$\delta W = \boldsymbol{B} \cdot d\boldsymbol{M}$$

**Proposition**: For a magnetizable system in a magnetic field, the free energy obeys

$$dF = -S\,dT - p\,dV - \boldsymbol{B} \cdot d\boldsymbol{M} + \mu\,dN$$

Proof: $dU = \delta Q - \delta W = T\,dS - \boldsymbol{B} \cdot d\boldsymbol{M}$ for purely magnetic work. The rest follows as before.

***Remark 1:*** By means of a Legendre transformation, $\tilde{F} = F + \boldsymbol{B} \cdot \boldsymbol{M}$, one can switch from $F = F(T, V, M, N)$ to $\tilde{F} = \tilde{F}(T, V, B, N)$. Similar things can be done for the other thermodynamic potentials.

## 1.5 Quantum Statistical Mechanics

### 1.5.1 The Postulates of Quantum Statistical Mechanics

Let a system at time $t$ be in a quantum state $|\psi(t)\rangle$. Expand this state in a complete set of orthonormal functions $\{|\varphi_n\rangle\}$.

$$|\psi(t)\rangle = \sum_n C_n(t)\,|\varphi_n\rangle\,, \qquad C_n(t) \in \mathbb{C}$$

***Remark 1:*** $|C_n(t)|^2$ is the probability for a system to be in state $|\varphi_n\rangle$ at time $t$.

***Remark 2:*** The time evolution of $|\psi(t)\rangle$ is determined by Schrodinger's equation:

$$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = \hat{H}\,|\psi(t)\rangle$$

with $\hat{H}$ representing the systems <u>Hamiltonian operator</u>, and we assume $\hat{H}$ to be time-independent. A deterministic system would imply that if you know state $|\psi(0)\rangle$, then you know the system's state at all future times.

***Remark 3:***

$$\langle\varphi_m|\psi(t)\rangle = \sum_n C_n\,\langle\varphi_m|\varphi_n\rangle = C_m(t)$$

$$\begin{aligned}
\Rightarrow i\hbar\,\dot{C}_n(t) &= \left\langle\varphi_n\middle|i\hbar\frac{\partial}{\partial t}\psi(t)\right\rangle = \left\langle\varphi_n\middle|\hat{H}\middle|\psi(t)\right\rangle \\
&= \sum_m C_m(t)\left\langle\varphi_n\middle|\hat{H}\middle|\varphi_m\right\rangle \\
&:= \sum_m H_{nm}C_m(t)
\end{aligned}$$

where $H_{nm}$ are the matrix elements of the Hamiltonian.

$$\begin{aligned}
\Rightarrow \dot{C}_n(t) &= \frac{i}{\hbar}\sum_m H_{nm}C_n(t) \\
\dot{C}_n^* &= \frac{i}{\hbar}\sum_m H_{nm}^*C_m^*(t) \\
&= \frac{i}{\hbar}\sum_m H_{mn}C_m^*(t)
\end{aligned}$$

Here $\langle\varphi|\psi\rangle$ is a scalar product in Hilbert space and we have used the fact that $\hat{H}$ is hermitian.

$$\hat{H}^\dagger = \hat{H} \quad\Rightarrow\quad H_{nm}^* = H_{mn}$$

Let $\hat{A}$ be an operator in Hilbert space that corresponds to some observable $\mathcal{A}$. From Phys 515/516, we know thath the expectation of $\hat{A}$ in a state $\psi(t)$ is given by:

$$\begin{aligned}
\left\langle\hat{A}\right\rangle &= \left\langle\psi(t)\middle|\hat{A}\middle|\psi(t)\right\rangle \\
&= \sum_{nm} C_n^*(t)C_m(t)\left\langle\varphi_n\middle|\hat{A}\middle|\varphi_m\right\rangle \\
&= \sum_{nm} C_n^*(t)C_m(t)A_{nm}
\end{aligned}$$

***Remark 4:*** The normalization of $|\psi(t)\rangle$ implies that

$$\sum_n |C_n(t)|^2 = 1$$

***Remark 5:*** $\left\langle \hat{A} \right\rangle$ is the average value obtained for the observable, taken from a large number of simultaneous measurements on equally prepared systems.

***Remark 6:*** Quantum Mechanics implies an averaging concept even in the absence of statistical mechanics considerations.

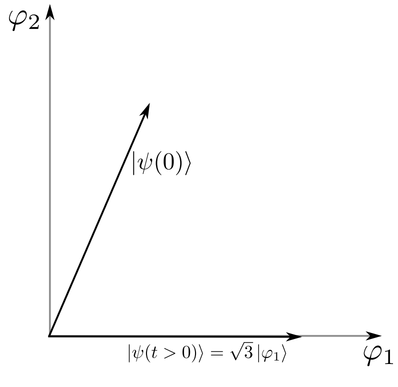
***Remark 7:*** Due to the butterfly effect, we are never quite sure which state the system is in, and therefore we need an ensemble average. Thus the ensemble averaged values of $\left\langle \hat{A} \right\rangle$ (which is now a random variable due to butterflies) is:

$$\langle A \rangle = \sum_{nm} \langle C_n^*(t) C_m(t) \rangle \, A_{nm}$$

For a deterministic quantum mechanical system, $|\psi(t)\rangle$ describes a system for all time given $|\psi(0)\rangle$. For example, let $|\psi(0)\rangle = |\varphi_1\rangle + 2\,|\varphi_2\rangle$:

$$\begin{aligned} \Rightarrow C_1(t=0) &= 1 \\ C_2(t=0) &= 2 \end{aligned}$$

In 2 dimensional Hilbert space:



The butterfly effect means we are not *exactly* sure what the state of $|\psi(0)\rangle$ is.

$$\Rightarrow \hat{H}_{total} = \hat{H}_{system} + \hat{H}_{system+butterflies}$$

**Definition 1:** The density matrix or statistical operator of the system is defined as the operator $\hat{\rho}$ whose matrix elements are

$$\rho_{nm}(t) = \langle \varphi_n | \hat{\rho} | \varphi_m \rangle := \langle C_m(t) C_n^*(t) \rangle$$

***Remark 8:*** As in §4.1, the average above can be considered either as a time average for one system, or as an ensemble average at a given time.

**Proposition 1**: Let $\hat{\rho}$ be the statistical operator of some system. Let $\hat{A}$ be an operator that corresponds to observable $\mathcal{A}$. Then the ensemble average $\left\langle \hat{A} \right\rangle$ of the expectation value $A$ of $\hat{A}$ is given by:

$$\langle A \rangle = \operatorname{tr}\left(\hat{\rho}\hat{A}\right)$$

Proof:

$$\begin{aligned}
\operatorname{tr}\left(\hat{\rho}\hat{A}\right) &= \sum_n \left\langle \varphi_n \middle| \hat{\rho}\hat{A} \middle| \varphi_n \right\rangle \\
&= \sum_n \sum_m \langle \varphi_n | \hat{\rho} | \varphi_m \rangle \left\langle \varphi_m \middle| \hat{A} \middle| \varphi_n \right\rangle \\
&= \sum_{nm} \rho_{nm} A_{mn} \\
&= \sum_{nm} \langle C_m(t) C_n^*(t) \rangle A_{mn} \\
&= \sum_{nm} \langle C_n(t) C_m^*(t) \rangle A_{nm} \\
&= \langle A \rangle
\end{aligned}$$

**Proposition 2**: The statistical operator obeys the Liouville - von Neumann equation:

$$i\hbar \partial_t \hat{\rho}(t) = \left[\hat{H}, \hat{\rho}(t)\right]$$

Proof:

$$\begin{aligned}
i\hbar \partial_t \rho_{nm} &= i\hbar \left( \left\langle \dot{C}_n(t) C_m^*(t) + C_n(t) \dot{C}_m^*(t) \right\rangle \right) \\
&= \sum_s \left( H_{ns} \langle C_s(t) C_m^*(t) \rangle - H_{sm} \langle C_n(t) C_m^*(t) \rangle \right) \\
&= \sum_s \left( H_{ns} \rho_{sm} - \rho_{ns} H_{sm} \right) \\
&= \left(\hat{H}\hat{\rho}\right)_{nm} - \left(\hat{\rho}\hat{H}\right)_{nm} \\
&= \left[\hat{H}, \hat{\rho}\right]_{nm}
\end{aligned}$$

**Corollary 1**: For a system in equilibrium, the density matrix commutes with the Hamiltonian, $\left[\hat{H}, \hat{\rho}\right] = 0$.

Proof: In equilibrium, $\hat{\rho}$ is *not* a function of time.

$$\Rightarrow \partial_t \hat{\rho}_{eq} = 0 \quad \Rightarrow \quad \left[\hat{H}, \hat{\rho}\right] = 0$$

**Corollary 2**: In equilibrium, $\hat{\rho}$, can depend only on the Hamiltonian and operators that commute with the Hamiltonian. (I.e. that correspond to conserved quantities.)

Let $\varphi_n$ be energy eigenstates:

$$\hat{H}\varphi_n = E_n \varphi_n$$

Quantum mechanics tell us that, if $\hat{A}$ commutes with $\hat{H}$, then $\varphi_n$ can be chosen so that they are eigenstates of both $\hat{A}$ and $\hat{H}$.

**Corollary 3**: In this basis, the density operator is diagonal:

$$\rho_{nm} = \delta_{nm} \rho_n$$

where $\rho_n$ is a function of the $E_n$ and of the eigenvalues of any other conserved quantities that $\hat{\rho}$ depends on.

**Corollary 4**:

$$\text{Liouville's Equation:} \qquad \frac{d}{dt}\rho\left(p, q, t\right) = 0$$

**Postulate 1 - Postulate of Equal a priori Probabilities**: For an isolated system in equilibrium:

$$\rho_n = \rho(E_n)$$

with $\rho_n$ given by equation ($\star$) in §4.2.

***Remark 9:*** "Isolated" means isolated up to the butterfly effect.

**Postulate 2 - Approach to Equilibrium**: This is the exact same as done in §1.3.2.P2.

***Remark 10:*** Remaining job: to find the statistical operator for systems that are *not* isolated.

### 1.5.2 The Density Matrix for Various Ensembles

**Microcanonical Ensemble**: By §5.1, the statistical operator in terms of its matrix elements in a basis of energy eigenstates is:

$$\rho_n = \frac{1}{\sum 1} cdot \begin{cases} 1, & E - \Delta E < E_n < E \\ 0, & \text{otherwise} \end{cases}$$

**Canonical Ensemble**: By §4.3, the correct matrix elements are given by:

$$\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{H}}$$

with $Z = \operatorname{tr}\left(e^{-\beta \hat{H}}\right)$ being the canonical partition function. We can also write

$$\begin{aligned} Z &= \operatorname{tr}\left(e^{-\beta \hat{H}}\right) \\ &= \sum n \left\langle \varphi_n \middle| e^{-\beta \hat{H}} \middle| \varphi_n \right\rangle \\ &= \sum_n \left\langle \varphi_n \middle| e^{-\beta E_n} \middle| \varphi_n \right\rangle \\ &= \sum_n e^{-\beta E_n} \end{aligned}$$

**Grand Canonical Ensemble**: By §1.4.4:

$$\hat{\rho} = \frac{1}{Z} e^{-\beta\left(\hat{H} - \mu \hat{N}\right)}$$

where $\hat{N}$ is the number operator and

$$Z = \operatorname{tr}\left(e^{-\beta\left(\hat{H} - \mu \hat{N}\right)}\right)$$

is the grand partition function. We know that

$$\hat{H} \left|\varphi_{n,N}\right\rangle = E_n \left|\varphi_{n,N}\right\rangle, \qquad \hat{N} \left|\varphi_{n,N}\right\rangle = N_n \left|\varphi_{n,N}\right\rangle$$

And thus we can write $Z$ as:

$$\begin{aligned} Z &= \operatorname{tr}\left(e^{-\beta\left(\hat{H} - \mu \hat{N}\right)}\right) \\ &= \sum_{n,N} \left\langle \varphi_{n,N} \middle| e^{-\beta\left(\hat{H} - \mu \hat{N}\right)} \middle| \varphi_{n,N} \right\rangle \\ &= \sum_{n,N} \left\langle \varphi_{n,N} \middle| e^{-\beta(E_n - \mu N_n)} \middle| \varphi_{n,N} \right\rangle \quad \text{since } \left[\hat{H}, \hat{N}\right] = 0 \\ &= \sum_{n,N} e^{-\beta(E - n - \mu N_n)} \end{aligned}$$

***Remark 1:*** All thermodynamic quantities can be calculated from the partition function as before. In particular, in the canonical ensemble, we have

$$F = -k_B T \ln Z = -k_B T \ln \left[ \operatorname{tr} \left( e^{-\beta \hat{H}} \right) \right]$$

and in the grand canonical ensemble:

$$J = -k_B T \ln Z = -k_B T \ln \left[ \operatorname{tr} \left( e^{-\beta \left( \hat{H} - \mu \hat{N} \right)} \right) \right]$$

***Remark 2:*** Statistical mechanics, quantum or classical, then boils down to calculating partition functions.

### 1.5.3 Fermions and Bosons

Consider a quantum system consisting of $N$ particles, which, for simplicity, we will take to be non-interacting. Let the single particle energy levels be denoted $\mathcal{E}(n = 0, 1, 2, \ldots)$. Each of the corresponding eigenstates can be occupied by $0, 1, 2, \ldots$ particles.

**Postulate 1**: Quantum particles come in two species: fermions and bosons. For fermions, at most one particle can occupy a given energy eigenstate, while any number of bosons can occupy any given energy eigenstate.

***Remark 1:*** The restriction on fermions applies only to fermions of the same kind (eg, all electrons) whose quantum numbers are the same.

***Remark 2:*** The fermionic restriction is known as the Pauli Exclusion Principle.

***Remark 3:*** Fermions have half-integer spin, whereas bosons have integer spin.

¡++¿ ¡++¿