Handout 5. Microcanonical Ensemble

January 19, 2011

Contents

1 Properties of flow in phase space
   1.1 Trajectories in phase space ........................................... 2
   1.2 One trajectory over long time ............................................ 3
   1.3 An ensemble of points flowing in phase space ....................... 6

2 Microcanonical Ensemble
   2.1 Uniform density assumption ............................................. 7
   2.2 Ideal Gas .................................................................. 9
   2.3 Entropy .................................................................. 13
The purpose of this lecture is

1. To justify the “uniform” probability assumption in the microcanonical ensemble.
2. To derive the momentum distribution of one particle in an ideal gas (in a container).
3. To obtain the entropy expression in microcanonical ensemble, using ideal gas as an example.

Reading Assignment: Sethna § 3.1, § 3.2.

1 Properties of flow in phase space

1.1 Trajectories in phase space

Q: What can we say about the trajectories in phase space based on classical mechanics?

A:

1. Flow line (trajectory) is completely deterministic

\[
\begin{align*}
\dot{q}_i &= \frac{\partial H}{\partial p_i} \\
\dot{p}_i &= -\frac{\partial H}{\partial q_i}
\end{align*}
\] (1)

Hence two trajectories never cross in phase space.

This should never happen, otherwise the flow direction of point \( P \) is not determined.

2. Liouville’s theorem

\[
\frac{d\rho}{dt} \equiv \frac{\partial \rho}{\partial t} + \{\rho, H\} = 0
\] (2)

So there are no attractors in phase space.
This should never happen, otherwise $d\rho/dt > 0$. Attractor is the place where many trajectories will converge to. The local density will increase as a set of trajectories converge to an attractor.

3. Consider a little “cube” in phase space. (You can imagine many copies of the system with very similar initial conditions. The cube is formed all the points representing the initial conditions in phase space.) Let the initial density of the points in the cube be uniformly distributed. This can be represented by a density field $\rho(\mu, t = 0)$ that is uniform inside the cube and zero outside.

As every point inside the cube flows to a different location at a later time, the cube is transformed to a different shape at a different location.

Due to Liouville’s theorem, the density $\rho$ remain $\rho = c$ (the same constant) inside the new shape and $\rho = 0$ outside. Hence the volume of the new shape remains constant ($V_0$).

1.2 One trajectory over long time

Q: Can a trajectory from on point $\mu_1$ in phase space always reach any other point $\mu_2$, given sufficiently long time?

A: We can imagine the following possibilities:
1. We know that the Hamiltonian $H$ (total energy) is conserved along a trajectory.

So, there is no hope for a trajectory to link $\mu_1$ and $\mu_2$ if $H(\mu_1) \neq H(\mu_2)$.

Hence, in the following discussion we will assume $H(\mu_1) = H(\mu_2)$, i.e. $\mu_1$ and $\mu_2$ lie on the same constant energy surface: $H(\mu) = E$. This is a $(6N - 1)$-dimensional hyper-surface in the $6N$-dimensional phase space.

2. The trajectory may form a loop. Then the trajectory will never reach $\mu_2$ if $\mu_2$ is not in the loop.

3. The constant energy surface may break into several disjoint regions. If $\mu_1$ and $\mu_2$ are in different regions, a trajectory originated from $\mu_1$ will never reach $\mu_2$.

Example: pendulum

4. Suppose the trajectory does not form a loop and the constant energy surface is one continuous surface. The constant energy surface may still separate into regions where trajectories originated from one region never visit the other region.

— This type of system is called non-ergodic.
5. If **none** of the above (1-4) is happening, the system is called Ergodic. In an ergodic system, we still cannot guarantee that a trajectory starting from $\mu_1$ will exactly go through any other part $\mu_2$ in phase space. (This is because the dimension of the phase space is so high, hence there are too many points in the phase space. One trajectory, no matter how long, is a one-dimensional object, and can “get lost” in the phase space, i.e. not “dense enough” to sample all points in the phase space.)

But the trajectory can get arbitrarily close to $\mu_2$.

“At time $t_1$, the trajectory can pass by the neighborhood of $\mu_2$. At a later time $t_2$, the trajectory passes by $\mu_2$ at an even smaller distance...”

After a sufficiently long time, a single trajectory will visit the neighborhood of every point in the constant energy surface.

— This is the property of an ergodic system. Ergodicity is ultimately an assumption, because mathematically it is very difficult to prove that a given system (specified by its Hamiltonian) is ergodic.
1.3 An ensemble of points flowing in phase space

Now imagine a small cube (page 1) contained between two constant-energy surfaces $H = E$, $H = E + \Delta E$.

As all points in the cube flows in phase space. The cube transforms into a different shape but its volume remains $V_0$.

The trajectories of many non-linear systems with many degrees of freedom is chaotic, i.e. two trajectories with very similar initial conditions will diverge exponentially with time.

Q. How can the volume $V_0$ remain constant while all points in the original cube will have to be very far apart from each other as time increases?

A: The shape of $V_0$ will become very complex, e.g. it may consists of many thin fibers distributed almost every where between the two constant energy surface.

At a very late time $t$, $\rho(\mu, t)$ still has the form of

$$\rho(\mu, t) = \begin{cases} C & \mu \in V_0 \\ 0 & \mu \notin V_0 \end{cases} \quad (3)$$

except that the shape of $V_0$ is distributed almost every where between constant energy surface.
A density function $\rho(\mu, t)$ corresponds to an ensemble of points in phase space. Suppose we have a function $A(\mu)$ defined in phase space. (In the pendulum example, we have considered $A = \theta^2$.)

The average of function $A(\mu)$ over all points in the ensemble is called the **ensemble average**.

If $\rho$ changes with time, then the ensemble average is time dependent

$$\langle A \rangle(t) \equiv \int d^{6N} \mu A(\mu) \rho(\mu, t)$$

From experience, we know that many systems will reach an equilibrium state if left alone for a long time. Hence we expect the following limit to exist:

$$\lim_{t \to \infty} \langle A \rangle(t) = \langle A \rangle_{eq}$$

$\langle A \rangle_{eq}$ is the “equilibrium” ensemble average.

Q: Does this mean that

$$\lim_{t \to \infty} \rho(\mu, t) = \rho_{eq}(\mu)?$$

No. In previous example, no matter how large $t$ is,

$$\rho(\mu, t) = \begin{cases} C & \mu \in V_0 \\ 0 & \mu \notin V_0 \end{cases}$$

The only thing that changes with $t$ is the shape of $V_0$. The shape continues to transform with time, becoming thinner and thinner but visiting the neighborhood of more and more points in phase space.

So, $\lim_{t \to \infty} \rho(\mu, t)$ **DOES NOT EXIST**!

What’s going on?

## 2 Microcanonical Ensemble

### 2.1 Uniform density assumption

In Statistical Mechanics, an ensemble (microcanonical ensemble, canonical ensemble, grand canonical ensemble, ...) usually refers to an **equilibrium** density distribution $\rho_{eq}(\mu)$ that does not change with time.

The macroscopically measurable quantities is **assumed** to be an ensemble average over $\rho_{eq}(\mu)$.

$$\langle A \rangle_{eq} \equiv \int d^{6N} \mu A(\mu) \rho_{eq}(\mu)$$
In the **microcanonical ensemble**, we assume $\rho_{eq}$ to be uniform inside the entire region between the two constant energy surfaces, i.e.

$$
\rho_{eq}(\mu) = \rho_{mc}(\mu) = \begin{cases} 
C' & E \leq H(\mu) \leq E + \Delta E \\
0 & \text{otherwise}
\end{cases}
$$

There is nothing “micro” in the microcanonical ensemble. It’s just a name with an obscure historical origin.

Q: How do we justify the validity of the microcanonical ensemble assumption, given that $\lim_{t \to \infty} \rho(\mu, t) \neq \rho_{mc}(\mu)$ (recall previous section)?

A:

1. As $t$ increases, $\rho(\mu, t)$ becomes a highly oscillatory function changing volume rapidly between $C$ and $0$, depending on whether $\mu$ is inside volume $V_0$ or not.

   ![Diagram](image1)

   But if function $A(\mu)$ is smooth function, as is usually the case, then it is reasonable to expect

   $$
   \lim_{t \to \infty} \int d^6N \mu A(\mu) \rho(\mu, t) = \int d^6N \mu A(\mu) \rho_{mc}(\mu)
   $$

   In other words, $\lim_{t \to \infty} \rho(\mu, t)$ and $\rho_{eq}(\mu)$ give the same ensemble averages.

2. A reasonable assumption for $\rho_{eq}(\mu)$ must be time stationary, i.e.

   $$
   \frac{\partial \rho_{eq}}{\partial t} = -\{\rho_{eq}, H\} = 0
   $$

   Notice that

   $$
   \rho_{mc}(\mu) = [\Theta(H(\mu) - E) - \Theta(H(\mu) - E - \Delta E)] \cdot C'
   $$

   where $\Theta(x)$ is the step function.
Because $\rho_{mc}$ is a function of $H \Rightarrow \{\rho_{mc}, H\} = 0$.

Hence

$$\frac{\partial \rho_{mc}}{\partial t} = 0$$

The microcanonical ensemble distribution $\rho_{mc}$ is stationary!

3. The microcanonical ensemble assumption is consistent with the subjective probability assignment. If all we know about the system is that its total energy $H$ (which should be conserved) is somewhere between $E$ and $E + \Delta E$, then we would like to assign equal probability to all microscopic microstate $\mu$ that is consistent with the constraint $E \leq H(\mu) \leq E + \Delta E$.

2.2 Ideal Gas

Ideal gas is an important model in statistical mechanics and thermodynamics. It refers to $N$ molecules in a container. The interaction between the particles is sufficiently weak so that it will be ignored in many calculations. But conceptually, the interaction cannot be exactly zero, otherwise the system would no longer be ergodic — a particle would never be able to transfer energy to another particle and to reach equilibrium when there were no interactions at all.

Consider an ensemble of gas containers containing ideal gas particles (monatomic molecules) that can be described by the microcanonical ensemble.

Q: What is the velocity distribution of on gas particle?

The Hamiltonian of $N$-ideal gas molecules:

$$H(\{q_i\}, \{p_i\}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} + \sum_{i=1}^{N} \phi(x_i)$$
where $\phi(x)$ is the potential function to represent the effect of the gas container

$$
\phi(x) = \begin{cases} 
0 & \text{if } x \in V \text{ (volume of the container)} \\
\infty & \text{if } x \notin V 
\end{cases} 
$$

(15)

This basically means that $x_i$ has to stay within volume $V$ and when this is the case, we can neglect the potential energy completely.

$$
H(\{q_i\}, \{p_i\}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} 
$$

(16)

The constant energy surface $H(\{q_i\}, \{p_i\}) = E$ is a sphere in $3N$-dimensional space, i.e.,

$$
\sum_{i=1}^{3N} p_i^2 = 2mE = R^2 
$$

(17)

with radius $R = \sqrt{2mE}$.

Let’s first figure out the constant $C'$ in the microcanonical ensemble,

$$
\rho_{mc}(\mu) = \begin{cases} 
C' & E \leq H(\mu) \leq E + \Delta E \\
0 & \text{otherwise} 
\end{cases} 
$$

(18)

Normalization condition:

$$
1 = \int d^{6N} \mu \rho_{mc}(\mu) = \int_{E \leq H(\mu) \leq E + \Delta E} d^{6N} \mu C' = \left[ \tilde{\Omega}(E + \Delta E) - \tilde{\Omega}(E) \right] \cdot C' 
$$

(19)

where $\tilde{\Omega}(E)$ is the phase space volume of region $H(\mu) \leq E$ and $\tilde{\Omega}(E + \Delta E)$ is the phase space volume of region $H(\mu) \leq E + \Delta E$. This leads to

$$
C' = \frac{1}{\tilde{\Omega}(E + \Delta E) - \tilde{\Omega}(E)} 
$$

(20)

How big is $\tilde{\Omega}(E)$?

$$
\tilde{\Omega}(E) = \int_{H(\mu) \leq E} d^{6N} \mu = V^N \cdot \int_{\sum_{i=1}^{3N} p_i^2 \leq 2mE} dp_1 \cdots dp_N 
$$

(21)

Here we need to invoke an important mathematical formula. The volume of a sphere of radius $R$ in $d$-dimensional space is$^1$

$$
\mathcal{V}_{sp}(R,d) = \frac{\pi^{d/2} R^d}{(d/2)!} 
$$

(22)

$^1$It may seem strange to have the factorial of a half-integer, i.e. $(d/2)!$. The mathematically rigorous expression here is $\Gamma(d/2 + 1)$, where $\Gamma(x)$ is the Gamma function. It is defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. When $x$ is a positive integer, $\Gamma(x) = (x-1)!$. When $x$ is not an integer, we still have $\Gamma(x+1) = x \Gamma(x)$. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Hence $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2} \sqrt{\pi}$, $\Gamma\left(\frac{5}{2}\right) = \frac{3}{4} \sqrt{\pi}$, etc. We can easily verify that $\mathcal{V}_{sp}(R,3) = \frac{4}{3} \pi R^3$ and $\mathcal{V}_{sp}(R,2) = \pi R^2$. 

The term behind \( V^N \) is the volume of a sphere of radius \( R = \sqrt{2mE} \) in \( d = 3N \) dimensional space. Hence,

\[
\tilde{\Omega}(E) = V^N \cdot \frac{\pi^{3N/2} R^{3N}}{(3N/2)!} \tag{23}
\]

\[
\lim_{\Delta E \to 0} \frac{\tilde{\Omega}(E + \Delta E) - \tilde{\Omega}(E)}{\Delta E} = \frac{\partial \tilde{\Omega}(E)}{\partial E} = \frac{3N}{2} \frac{1}{E} \frac{(2\pi mE)^{3N/2}}{(3N/2)!} V^N = \frac{(2\pi m)^{3N/2} E^{3N/2-1} V^N}{(3N/2 - 1)!} \tag{24}
\]

In the limit of \( \Delta E \to 0 \), we can write

\[
C' = \frac{1}{\Delta E} \cdot \frac{\Delta E}{\tilde{\Omega}(E + \Delta E) - \tilde{\Omega}(E)} = \frac{1}{\Delta E} \cdot \frac{(3N/2 - 1)!}{(2\pi m)^{3N/2} E^{3N/2-1} V^N} \tag{25}
\]

Q: What is the probability distribution of \( p_1 \) — the momentum of molecule \( i = 1 \) in the \( x \)-direction?

A: The probability distribution function for \( p_1 \) is obtained by integrating the joint distribution function \( \rho_{mc}(q_1, \cdots, q_{3N}, p_1, \cdots, p_{3N}) \) over all the variables except \( p_1 \).

\[
f(p_1) = \int dp_2 \cdots dp_{3N} \cdot dq_1 dq_2 \cdots dq_{3N} \rho_{mc}(q_1, \cdots, q_{3N}, p_1, \cdots, p_{3N})
\]

\[
= \int_{2mE \leq \sum_{i=1}^{3N} p_i^2 \leq 2m(E+\Delta E)} dp_2 \cdots dp_{3N} V^N C'
\]

\[
= \int_{2mE - p_1^2 \leq \sum_{i=2}^{3N} p_i^2 \leq 2m(E+\Delta E) - p_1^2} dp_2 \cdots dp_{3N} V^N C'
\]

\[
= \left[ V_{sp} \left( \sqrt{2m(E+\Delta E) - p_1^2}, 3N - 1 \right) - V_{sp} \left( \sqrt{2mE - p_1^2}, 3N - 1 \right) \right] V^N C' \tag{26}
\]
In the limit of $\Delta E \to 0$,

\[
\frac{\mathcal{V}_{\text{sp}} \left( \sqrt{2m(E + \Delta E) - p_1^2}, 3N - 1 \right) - \mathcal{V}_{\text{sp}} \left( \sqrt{2mE - p_1^2}, 3N - 1 \right)}{\Delta E} = \frac{\partial}{\partial E} \mathcal{V}_{\text{sp}} \left( \sqrt{2mE - p_1^2}, 3N - 1 \right)
\]

\[
= \frac{\partial}{\partial E} \left[ \pi^{(3N-1)/2} \frac{(2mE - p_1^2)^{(3N-1)/2}}{(3N-1)!} \right]
\]

\[
= \frac{3N - 1}{2} \frac{2m}{2mE - p_1^2} \pi^{(3N-1)/2} \frac{(2mE - p_1^2)^{(3N-1)/2}}{(3N-1)!}
\]

\[
= 2m \frac{\pi^{(3N-1)/2} (2mE - p_1^2)^{(3N-1)/2}}{(3N-1)!}
\]

(27)

Returning to $f(p_1)$, and only keep the terms that depend on $p_1$,

\[
f(p_1) \propto (2mE - p_1^2)^{(3N-1)/2} \propto \left( 1 - \frac{p_1^2}{2mE} \right)^{(3N-1)/2}
\]

(28)

Notice the identity

\[
\lim_{n \to \infty} \left( 1 + \frac{x}{n} \right)^n = e^x
\]

(29)

and that $N \approx N - 1$ in the limit of large $N$. Hence, as $N \to \infty$,

\[
f(p_1) \propto \left( 1 - \frac{2}{3N} \frac{3N}{2E} \frac{p_1^2}{2m} \right)^{3N/2} \exp \left( -\frac{p_1^2}{2m} \frac{3N}{2E} \right)
\]

(30)

Using the normalization condition

\[
\int_{-\infty}^{\infty} dp_1 f(p_1) = 1
\]

(31)

we have,

\[
f(p_1) = \frac{1}{\sqrt{2\pi m(2E/3N)}} \exp \left( -\frac{p_1^2}{2m} \frac{3N}{2E} \right)
\]

(32)

Later on we will show that for an ideal gas (in the limit of large $N$),

\[
E = \frac{3}{2} N k_B T
\]

(33)

where $T$ is temperature and $k_B$ is Boltzmann’s constant. Hence

\[
f(p_1) = \frac{1}{\sqrt{2\pi mk_B T}} \exp \left( -\frac{p_1^2}{2m} \frac{3N}{k_B T} \right)
\]

(34)

Notice that $p_1^2/2m$ is the kinetic energy associated with $p_1$. Hence $f(p_1)$ is equivalent to Boltzmann’s distribution that will be derived later (in canonical ensemble).
2.3 Entropy

Entropy is a key concept in both thermodynamics and statistical mechanics, as well as in information theory (a measure of uncertainty or lack of information). In information theory, if an experiment has $N$ possible outcomes with equal probability, then the entropy is

$$S = k_B \log N$$  \hspace{1cm} (35)

In microcanonical ensemble,

$$S(N, V, E) = k_B \log \left[ \frac{\text{number of microscopic states between the constant energy surfaces:}}{E \leq H(\mu) \leq E + \Delta E} \right]$$  \hspace{1cm} (36)

For an ideal gas,

$$S(N, V, E) = k_B \log \frac{\tilde{\Omega}(E + \Delta E) - \tilde{\Omega}(E)}{N! h^{3N}}$$  \hspace{1cm} (37)

The numerator inside the log is the volume of the phase space between the two constant energy surfaces. $h$ is Planck’s constant, which is the fundamental constant from quantum mechanics.

Yes, even though we only discuss classical equilibrium statistical mechanics, a bare minimum of quantum mechanical concepts is required to fix some problems in classical mechanics.

We can view this as another evidence that classical mechanics is really just an approximation and quantum mechanics is a more accurate description of our physical world. Fortunately, these two terms can be intuitively understandable without working with quantum mechanics equations. The following are the justifications of the two terms in the denominator.

1. $N!$ term: Quantum mechanics says that the gas molecules are all identical or indistinguishable. Even though we would like to label molecules as $1, 2, ..., N$ there is really no way for us to tell which one is which! Therefore, two molecular structures with coordinates: $x_1 = (1, 2, 3)$, $x_2 = (4, 5, 6)$ and $x_1 = (4, 5, 6)$, $x_2 = (1, 2, 3)$ are indistinguishable from each other.

Swapping the location between two molecules does not give a new microscopic state.

2. $h^{3N}$ term: $h = 6.626 \times 10^{-34}$ J·s is Planck’s constant.

The numerator, $\tilde{\Omega}(E)$, is the phase space volume and has the unit of (momentum · distance)$^{3N}$.

The term inside log has to be dimensionless, otherwise, the magnitude of entropy would depend on our choices for the units of length, time, mass, and etc, which would be clearly absurd.

$h$ has the unit of momentum · distance. Therefore $h^{3N}$ has exactly the right unit to make the entire term inside the log dimensionless.
The uncertainty principle in quantum mechanics states that we cannot measure both the position and the momentum of any particle to infinite accuracy. Instead, their error bar must satisfy the relation:

$$\Delta q_i \cdot \Delta p_i \geq \hbar$$

for any \(i = 1, \ldots, 3N\) \hspace{1cm} (38)

Therefore, \(\tilde{\Omega}(E)/(N! \hbar^{3N})\) gives us the number of distinguishable states contained inside a phase space volume of \(\Omega(E)\).

We can show that the entropy expression for the ideal gas in microcanonical ensemble is

$$S(N, V, E) = Nk_B \left[ \log \left( \frac{V}{N} \left( \frac{4\pi mE}{3N\hbar^2} \right)^{3/2} \right) + \frac{5}{2} \right]$$

(Sackur-Tetrode formula) \hspace{1cm} (39)

We will derive the Sackur-Tetrode formula later. (Stirling’s formula is used to derive it.)

Define number density \(\rho = \frac{N}{V}\), and de Broglie wavelength

$$\lambda = \frac{\hbar}{\sqrt{4\pi mE/3N}} \left( = \frac{\hbar}{\sqrt{2\pi mk_B T}} \right)$$

(40)

then

$$S(N, V, E) = Nk_B \left[ \frac{5}{2} - \log(\rho\lambda^3) \right]$$

(41)

In molecular simulations, the **microcanonical ensemble** is usually referred to as the **NVE ensemble**.