2.1 Probabilities

In the previous lectures the theory of thermodynamics was formulated as a purely phenomenological theory, based on the (four) laws of thermodynamics. The kinetic theory makes an attempt to derive these laws from first principle. Since the theory is based entirely on a probabilistic description of the dynamics, it is a good place to review some of the most basic properties of the probability theory.

2.1.1 Probability space

Consider a measurable space $(\Omega, \mathcal{F})$, where

- $\Omega$ is a sample space or a set of possible outcomes. For example, discrete $\Omega = \{\text{Head, Tail}\}$, continues $\Omega = \mathbb{R}$ or functional $\Omega = \mathbb{R}^R$.
- $\mathcal{F} \subseteq 2^\Omega$ is a $\sigma$-algebra or a set of subsets of $\Omega$ (i.e. $\mathcal{F} \subseteq 2^\Omega$), such that
  
  1. $\emptyset \in \mathcal{F}$
  2. if $A \in \mathcal{F}$ then $(\Omega - A) \in \mathcal{F}$
  3. if $A_i \in \mathcal{F}$ then $\cup_i A_i \in \mathcal{F}$

**Problem:** Show that if $A_i \in \mathcal{F}$ then $\cap_i A_i \in \mathcal{F}$.

**Solution:** $A_i \in \mathcal{F} \Rightarrow \cup_i A_i \in \mathcal{F} \Rightarrow \Omega - \cup_i A_i = \cap_i A_i \in \mathcal{F}$.

Then $(\Omega, \mathcal{F}, P)$ is called a probability space if $P : \mathcal{F} \to \mathbb{R}$ is a function satisfying the following axioms:
• **Positivity Axiom:** $\mathcal{P}(A) \geq 0$ for $A \in \mathcal{F}$.

• **Unitarity Axiom:** $\mathcal{P}(\Omega) = 1$.

• **Additivity Axiom:** if $A_i \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$ (i.e. pairwise disjoint sets) then $\mathcal{P}(\bigcup_i A_i) = \sum_i \mathcal{P}(A_i)$.

These are the famous Kolmogorov’s axioms which are necessary for constructing a mathematical theory of probabilities. From the physical viewpoint there is still work to be done. After centuries of debate it finally became clear that there are two consistent ways to assign probabilities to physical events. These are the objective and subjective interpretations of probabilities.

### 2.1.2 Interpretation of probabilities

Objective (or experimental) probabilities are assigned by relative frequencies of observed events. For example, if we toss a fair coin $N$ times than roughly $N_{\text{Tail}} \approx N/2$ we will get tails. Then the objective probability of tails is defined as

$$\mathcal{P}_{\text{Experiment}}(\text{Tail}) = \frac{N_{\text{Tail}}}{N} \approx \frac{1}{2}$$

(2.1)

where $N$ is sufficiently large. Note that $N$ can never be infinite since we only have a finite time to do the experiment.

Subjective (or theoretical) probabilities are assigned according to degrees of belief in certain outcomes. The belief can (but does not have to) be based on what was previously observed. For example, even before we start tossing a fair coin we can examine it to determine that it “looks pretty symmetric” and thus the odds for both outcomes Tail and Head should be equal, i.e.

$$\mathcal{P}_{\text{Theory}}(\text{Tail}) = \frac{1}{2}$$

(2.2)

Some physicists claim that it is sufficient to only use objective probabilities, when in fact an agreement between subjective and objective probabilities is what distinguished a good theory from a bad theory. For example, it is often argued that cosmological theories cannot be tested subjectively since we can only make one observation, i.e. $N = 1$ in contrast to experiments in high energy physics where the number of observed similar events can be much larger $N \gg 1$. Of course, it is not $N$ what determines the success of a given theory, but an agreement of subjective probabilities with objective probabilities, i.e.

$$\mathcal{P}_{\text{Theory}} \approx \mathcal{P}_{\text{Experiment}}$$

(2.3)
The two numbers are never in a perfect agreement (or it would have been really surprising by itself) with each other.

Einstein’s quote: “In theory, theory and practice is the same. In practice, they are not.”

2.1.3 Random variables

In physics in general and in kinetic theory in particular, the random variables $x$ usually take continues values on a real line. Then the relevant probability space $(\Omega, \mathcal{F}, \mathcal{P})$ is given by

1. $\Omega = \mathbb{R}$ (or $\mathbb{R}^+$)
2. $\mathcal{F} = \{\text{finite or countable union of open intervals and their complements}\}$
3. $\mathcal{P}(a, b) = \int_a^b p(x)dx$ where $p(x)$ is a probability distribution.

Unlike the probability function $\mathcal{P} \leq 1$, the probability distribution function (PDF or probability distribution or simply distribution) $p \geq 0$ is unbounded from above.

*Expectation value* of an arbitrary function of random variable $f(x)$ is then defined as

$$\langle f \rangle \equiv \int_{-\infty}^{\infty} f(x)p(x)dx. \quad (2.4)$$

*Statistical moments* of the distribution are given by expectation values of powers of $x$, i.e.

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n p(x)dx. \quad (2.5)$$

*Characteristic function* is a Fourier transform of the distribution

$$\tilde{p}(k) = \langle e^{-ikx} \rangle = \int_{-\infty}^{\infty} e^{-ikx} p(x)dx. \quad (2.6)$$

which can be expressed in terms of statistical moments:

$$\tilde{p}(k) = \left( \sum_{n=0}^{\infty} \frac{(-ikx)^n}{n!} \right) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle. \quad (2.7)$$

*Cumulant-generating* function is natural logarithm of characteristic function with expansion in terms of *cumulants*:

$$\log \tilde{p}(k) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle_c. \quad (2.8)$$
CHAPTER 2. KINETIC THEORY

Relations between cumulants and moments can be obtained from (2.7) and (2.8) by collecting the term with the same power of \( k \) in the expansion, i.e.

\[
\log \left( 1 + \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle \right) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle_c.
\]

or

\[
\sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \left( \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle \right)^m = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle_c.
\]

The first few cumulants are given

1. Mean (by collecting \( k \) terms)

\[
\langle x \rangle_c = \langle x \rangle
\]

2. Variance (by collecting \( k^2 \) terms)

\[
\langle x^2 \rangle_c = \langle x^2 \rangle - \langle x \rangle^2
\]

3. Skewness (by collecting \( k^3 \) terms)

\[
\langle x^3 \rangle_c = \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x^3 \rangle
\]

2.1.4 Useful distributions

The most studied (in physics) probability distribution is the Gaussian (or normal distribution):

\[
p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}.
\]

It is completely specified by only two parameters, mean \( \mu \) and variance \( \sigma \). Possible deviations from Gaussianities are usually described by skewness, but it is by no means a unique measure of non-Gaussianities. In fact there are infinitely many non-Gaussian distributions whose skewness is zero. It is a subject of active research in cosmology to look for non-Gaussianities signals in the Cosmic Microwave Background (CMB) radiation. If such signals could be detected we would be able to learn something about cosmic strings (very heavy and very thin objects), or about the cosmological initial conditions.

Another useful example is the (Dirac) \( \delta \)-function. Strictly speaking is not a function, but a generalized function or simply a distribution. As for any generalized function it only makes sense to talk about \( \delta \)-function when it is inside of an integral. It can be defined by the following expression
\[ \int_{-\infty}^{\infty} \delta(x-a)f(x)dx = f(a) \]

where \( f(x) \) is an arbitrary function. Sometimes it is convenient to express \( \delta \)-function as a derivative of the Heaviside step function, i.e. \( \delta(x) = \frac{d}{dx}H(x) \).

We know how to assign equal probabilities to a finite set of events, i.e. \( \Omega = \{X_1, X_2, ..., X_N\} \). We know how to assign equal probabilities to a continuum of states between 0 and 1, i.e. \( \Omega = (0, 1) \). What we do not know is how to assign equal probabilities to integers, i.e. \( \Omega = \mathbb{Z} \). What is the probability that a random positive integer is even?

### 2.2 Liouville’s theorem

As was already mentioned, there are two approaches to thermodynamics:

- phenomenological (or macroscopic) and
- fundamental (or microscopic).

In both approaches we make assumptions (low energy density, equilibrium, etc.) and hope that the nature will respect them at least in some limit. An active area of research is to explore other limits (high energy density, equilibrium) under other assumptions. Now that we have seen a phenomenological formulation of thermodynamics we will switch to a derivation of thermodynamics from a more fundamental viewpoint.

Consider again a large system with \( 3N \) degrees of freedom (d.o.f.). For example, \( 6.02214 \times 10^{23} \) molecules in a box. Microstate of the system can be described by a \( 6N \) dimensional vector \( (q_1, q_2, ..., q_{3N}, p_1, p_2, ..., p_{3N}) \) in phase space \( \Gamma \), where \( \{q_1, q_2, ..., q_{3N}\} \) are the position coordinates and \( \{p_1, p_2, ..., p_{3N}\} \) are the momentum coordinates. Trajectories in phase space described by Hamiltonian equations of motion:

\[
\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad (2.13)
\]

and

\[
\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad (2.14)
\]

where \( i = 1...3N \). Consider representative points in an infinitesimal \( 6N \)-dimensional hypercube with volume

\[
d\Gamma = dq_1...dq_{3N}dp_1...dp_{3N} \quad (2.15)
\]
CHAPTER 2. KINETIC THEORY

In time $\delta t$ the representative point move to new locations

\begin{align}
q'_i &= q_i + \dot{q}_i \delta t \quad \Rightarrow \quad dq'_i = dq_i + d\dot{q}_i dt \\
p'_i &= p_i + \dot{p}_i \delta t \quad \Rightarrow \quad dp'_i = dp_i + d\dot{p}_i dt
\end{align}

(2.16) (2.17)

and the volume becomes distorted

\[ d\Gamma' = \left( dq_1 + \left( \frac{\partial \dot{q}_1}{\partial q_1} dq_1 + \ldots + \frac{\partial \dot{q}_1}{\partial p_1} dp_1 + \ldots \right) dt \right) \ldots \left( dp_1 + \left( \frac{\partial \dot{p}_1}{\partial q_1} dq_1 + \ldots + \frac{\partial \dot{p}_1}{\partial p_1} dp_1 + \ldots \right) dt \right) \ldots \]

(2.18)

By expanding to the linear order in $dt$ we obtain

\[ \frac{d\Gamma' - d\Gamma}{dt} = \sum_{i=1}^{3N} \left( \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = \sum_{i=1}^{3N} \left( \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0 \]

(2.19)

by Hamiltonian equations of motion. Alternatively, the divergence of the velocity flow $\mathbf{v} = (\dot{q}_1, \ldots, \dot{q}_{3N}, \dot{p}_1, \ldots, \dot{p}_{3N})$ vanishes, i.e.

\[ \nabla \cdot \mathbf{v} = \left( \frac{\partial}{\partial q_1}, \ldots, \frac{\partial}{\partial p_{3N}}, \frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_{3N}} \right) \cdot (\dot{q}_1, \ldots, \dot{q}_{3N}, \dot{p}_1, \ldots, \dot{p}_{3N}) = \sum_{i=1}^{3N} \left( \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0. \]

(2.20)

This is the famous Liouville’s theorem which says that the flow in a phase space of Hamiltonian systems is incompressible, i.e. $d\Gamma = d\Gamma'$.

The Liouville’s theorem can also be expressed in a differential form. Let us define a multiparticle distribution function $\rho(q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N})$ over the phase space, where

\[ \rho(q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N})dq_1\ldots dq_{3N}dp_1\ldots dp_{3N} \]

(2.21)

is a probability to find a system in a state between $(q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N})$ and $(q_1 + dq_1, \ldots, q_{3N} + dq_{3N}, p_1 + dp_1, \ldots, p_{3N} + dp_{3N})$. Then the continuity equation

\[ \frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial q_i}(\dot{q}_i \rho) + \frac{\partial \rho}{\partial p_i}(\dot{p}_i \rho) = 0 \]

(2.22)

and the incompressibility of flow leads to the Liouville equation

\[ \frac{\partial \rho}{\partial t} = - \sum_{i=1}^{3N} \left( \frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = - \{\rho, H\} \]

(2.23)

where the Poisson brackets are defined as

\[ \{A, B\} = \sum_{i=1}^{3N} \left( \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right). \]

(2.24)
The ensemble averages are defined by

$$\langle \mathcal{O} \rangle = \int \mathcal{O}(q, p) \rho(q, p, t) dq dp$$

(2.25)

and their evolution

$$\frac{d\langle \mathcal{O} \rangle}{dt} = \int \mathcal{O}(q, p) \frac{\partial \rho(q, p, t)}{\partial t} dq dp = - \int \mathcal{O} \frac{3N}{q} \left( \frac{\partial \rho \partial H}{\partial q_i \partial p_i} - \frac{\partial \rho \partial H}{\partial p_i \partial q_i} \right) dq dp$$

(2.26)

Thus, for the ensemble average of arbitrary $\mathcal{O}$ not to change with time (or to be in equilibrium) we must have $\frac{\partial \rho}{\partial t} = - \{\rho_{eq}, H\}$.

For the distributions which are not in equilibrium very little is known. First of all there are two types of averages: ensemble average (2.25), and time average

$$\langle \mathcal{O} \rangle_t = \frac{1}{t} \int \mathcal{O}(q(t), p(t)) dt.$$  

(2.28)

Ergodicity is the assumption that the two averages are equivalent which was proved for only relatively simple systems, but is believed to hold for more complicated systems as well.

Of course, the exact specification of the state might not be possible in experiments with a large number of d.o.f. It is convenient instead to define a single particle correlation function,

$$f(q, p) = N \int \rho(q_1 = q, p_1 = p, q_2, p_2, ..., q_N, p_N) dq_2 dp_2 ... dq_N dp_N.$$  

(2.29)

The main idea in the kinetic theory is to make some assumptions about unobservable microscopic information (e.g. $q$’s and $p$’s), and to concentrate on only observable macroscopic information (e.g. $V$ and $P$). Thus, we make two simplifying assumptions:

1. **Molecular chaos assumption:**

$$\rho(q_1, ..., q_{3N}, p_1, ..., p_{3N}) = f(q_1, p_1) f(q_2, p_2) ... f(q_N, p_N)$$

(2.30)

where $f(q, p)$ is a single particle distribution function, where

$$f(q, p) dq dp$$

(2.31)
is a probability to find a particle in a state between \((q, p)\) and \((q + dq, p + dp)\).

2. **Separation of scales assumption:**

Size of Box \(\gg\) Resolution Scale \(\gg\) Inter-molecular Separation \(\gg\) Molecular Size

The main problem in a kinetic theory is to find \(f(q, p, t)\) for a given type of interaction and external force \(F\).

### 2.3 Binary collisions

1. **Without collisions:**
   
   (a) A neighborhood \(dqdp\) moves to \(dq'dp' = dqdp\) since for Hamiltonian systems the flow is incompressible.
   
   (b) A molecule at \((q, p, t)\) moves to \((q + \frac{P}{m} \delta t, p + F \delta t, t + \delta t)\) in time \(\delta t\), i.e.
   
   \[
   f(q + \frac{P}{m} \delta t, p + F \delta t, t + \delta t) dqdp' = f(q, p, t) dqdp \quad (2.32)
   
   or
   
   \[
   f(q + \frac{P}{m} \delta t, p + F \delta t, t + \delta t) = f(q, p, t) \quad (2.33)
   
2. **With collisions:**

   \[
   f(q + \frac{P}{m} \delta t, p + F \delta t, t + \delta t) = f(q, p, t) + \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \delta t \quad (2.34)
   
   where \(F\) is a conservative force.

By expanding the left hand site in powers of \(\delta t\) and keeping only linear terms we get the transport equation

\[
\frac{\partial}{\partial t} f(q, p, t) + \frac{p}{m} \cdot \frac{\partial}{\partial q} f(q, p, t) + F \cdot \frac{\partial}{\partial p} f(q, p, t) = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \quad (2.35)
\]

which is not very useful unless \(\left(\frac{\partial f}{\partial t}\right)_{\text{coll}}\) is specified. Of course, finding or modeling the collision term is the biggest challenge in the kinetic theory. In the simplest model one only takes into account binary collisions and assumes that the colliding particles are uncorrelated (i.e. molecular chaos assumption).

Consider an elastic collision of two spherically symmetric (spin-less) molecules with \(m_1, p_1\) and \(m_2, p_2\). After collision their respective momenta are \(p'_1\) and \(p'_2\). Then the following conservation laws apply:
CHAPTER 2. KINETIC THEORY

- Momentum conservation:
  \[ p_1 + p_2 = p'_1 + p'_2. \] (2.36)

- Energy conservation:
  \[ \frac{p_1^2}{m_1} + \frac{p_2^2}{m_2} = \frac{p_1'^2}{m_1} + \frac{p_2'^2}{m_2}. \] (2.37)

Define:

- Reduced mass
  \[ \mu \equiv \frac{m_1 m_2}{m_1 + m_2}. \] (2.38)

- Relative momentum
  \[ p \equiv \frac{m_2 p_1 - m_1 p_2}{m_1 + m_2} = \mu \left( \frac{p_1}{m_1} - \frac{p_2}{m_2} \right). \] (2.39)

Then the conservation of energy can be written as
  \[ |p| = |p'|, \] (2.40)
i.e. collision simply rotates relative momentum. Such rotations can be specified by a unit vector in spherical coordinates \((\theta, \phi)\), where

- inclination angle \(\theta\) between \(p\) and \(p'\)
- azimuthal angle \(\phi\) of \(p'\) about \(p\).

The dynamical aspect of the collisions are described by differential cross section \(|d\sigma/d\Omega|\) - the cross-sectional area which scatters particles into solid angle \(d\Omega\) around \(\Omega\). For scattering of two spheres of diameter \(D\) with impact parameter \(b\) the scattering angle is given by
  \[ \cos \left( \frac{\theta}{2} \right) = \frac{b}{D} \] (2.41)
and thus, the cross-sectional area is
  \[ d\sigma = bd\phi db = D \cos \left( \frac{\theta}{2} \right) D \sin \left( \frac{\theta}{2} \right) d\theta d\phi = \frac{D^2}{4} \sin \theta d\theta d\phi = \frac{D^2}{4} d\Omega \] (2.42)
and
  \[ \frac{d\sigma}{d\Omega} = \frac{D}{4} \] (2.43)
The total cross-section is given by

$$\sigma_{\text{total}} = \int \left| \frac{d\sigma}{d\Omega} \right| d\Omega$$

which

$$\sigma_{\text{total}} = \int \frac{D^2}{4} d\Omega = \pi D^2$$

for an elastic scattering of hard spheres.

In quantum mechanics the scattering probabilities $\mathcal{P}_{p_1,p_2 \rightarrow p'_1,p'_2}$ per unit time per unit volume are expressed through scattering amplitudes $A_{p_1,p_2 \rightarrow p'_1,p'_2}$ using Born rule,

$$\mathcal{P}_{p_1,p_2 \rightarrow p'_1,p'_2} = \delta_{(p_1+p_2-p'_1-p'_2)} \delta_{(p_1^2+p_2^2-p'_1^2-p'_2^2)} |A_{p_1,p_2 \rightarrow p'_1,p'_2}|^2$$

$$= \delta_{(p_1+p_2-p'_1-p'_2)} \delta_{(p_1^2+p_2^2-p'_1^2-p'_2^2)} |\langle p'_1,p'_2|T|p_1,p_2 \rangle|^2$$

where $T$ is the quantum mechanical transition matrix and all of the factors of $(2\pi)$ where absorbed into definition of $A_{p_1,p_2 \rightarrow p'_1,p'_2}$. If we assume that the interactions are invariant under spatial rotations, reflections and time reversal (e.g. electromagnetic interactions), then the transition matrix $T$ has the following symmetries

$$\langle p'_1,p'_2|T|p_1,p_2 \rangle = \langle Rp'_1,Rp'_2|T|Rp_1,Rp_2 \rangle$$

$$\langle p'_1,p'_2|T|p_1,p_2 \rangle = \langle -p_1,-p_2|T|p'_1,-p'_2 \rangle,$$

where $R$ is a matrix representing spatial rotations and/or reflections. Note that the vector $p$ can also include spin d.o.f., then $R$ rotates the spatial and spin d.o.f., but reflects only spatial d.o.f. It follows form (2.47) that

$$\langle p'_1,p'_2|T|p_1,p_2 \rangle = \langle p_1,p_2|T|p'_1,p'_2 \rangle.$$
One can apply the ideas of kinetic theory to other systems with a large number of d.o.f. For example, a very large network of cosmic strings can be described by a transport equation analogous to the Boltzmann transport equation for particles.

### 2.4 Boltzmann equation, H-theorem

#### 2.4.1 Boltzmann equation

The collision term \( \left( \frac{\partial f}{\partial t} \right)_{\text{coll}} \) in the transport equation

\[
\frac{\partial}{\partial t} f(q, p, t) + \frac{p}{m} \cdot \frac{\partial}{\partial q} f(q, p, t) + F \cdot \frac{\partial}{\partial p} f(q, p, t) = \left( \frac{\partial f}{\partial t} \right)_{\text{coll}}
\]  

(2.49)

can be expressed as

\[
\left( \frac{\partial f}{\partial t} \right)_{\text{coll}} = \bar{R} - R
\]  

(2.50)

where

- \( R \ dt \ dq \ dp \) is the number of collisions from time \( t \) to \( t + dt \) of particles with initial position from \( q \) to \( q + dq \) and initial momentum from \( p \) to \( p + dp \) and

- \( \bar{R} \ dt \ dq \ dp \) is the number of collisions from time \( t \) to \( t + dt \) of particles with final position from \( q \) to \( q + dq \) and final momentum from \( p \) to \( p + dp \).

With only binary collisions taken into account

\[
R(q, p_1) = \int dp_2 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} f(q, p_1) f(q, p_2)
\]  

(2.51)

and

\[
\bar{R}(q, p_1) = \int dp_2 dp'_2 P_{p'_1, p'_2 \rightarrow p_1, p_2} f(q, p'_1) f(q, p'_2)
\]  

(2.52)

where \( P_{p_1, p_2 \rightarrow p'_1, p'_2} \) is the probability density to go from initial state \( p_1, p_2 \) to final state \( p'_1, p'_2 \) in time \( dt \). It follows from symmetry considerations that \( P_{p_1, p_2 \rightarrow p'_1, p'_2} = P_{p'_1, p'_2 \rightarrow p_1, p_2} \). Thus,

\[
\left( \frac{\partial f}{\partial t} \right)_{\text{coll}} = \int dp_2 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} (f_1(p'_1) f(p'_2) - f(p_1) f(p_2))
\]  

(2.53)
CHAPTER 2. KINETIC THEORY

The rate of particles scattering per unit time into solid angle \( \Omega \) can be also expressed using differential crosses

\[
dp_1 dp_2 P_{p_1, p_2 \rightarrow p_1', p_2'} = |v_1 - v_2| \left( \frac{d\sigma}{d\Omega} \right) d\Omega \tag{2.54}
\]

then the collision term takes the following form

\[
\left( \frac{\partial f}{\partial t} \right)_{\text{coll}} = \int dp_2 |v_1 - v_2| \left( \frac{d\sigma}{d\Omega} \right) d\Omega (f(p_1') f(p_2') - f(p_1) f(p_2)) \tag{2.55}
\]

We can now plug in the collision term into (2.49) to obtain the famous Boltzmann transport equation:

\[
\frac{\partial}{\partial t} f(p_1) + \frac{p}{m} \cdot \frac{\partial}{\partial q} f(p_1) + F \cdot \frac{\partial}{\partial p} f(p_1) = \int dp_2 |v_1 - v_2| \left( \frac{d\sigma}{d\Omega} \right) d\Omega (f(p_1') f(p_2') - f(p_1) f(p_2)) \tag{2.56}
\]

This is only the first equation in a hierarchy of equations known as the BBGKY hierarchy. To obtain other equations one has to include many particle distributions between colliding particles.

2.4.2 H-theorem

It follows from Boltzmann’s transport equation that a local equilibrium is established when

\[
\frac{\partial}{\partial t} f(p_1) = \int dp_2 dp_1' dp_2' P_{p_1, p_2 \rightarrow p_1', p_2'} (f(p_1') f(p_2') - f(p_1) f(p_2)) = 0 \tag{2.58}
\]

Define the Boltzmann’s \( H \) function as

\[
H(t) \equiv \int dp f(p, t) \log f(p, t) \tag{2.59}
\]

then

\[
\frac{dH(t)}{dt} = \int dp \frac{\partial}{\partial t} [f(p, t) \log f(p, t)] = \int dp \frac{\partial f(p, t)}{\partial t} [1 + \log f(p, t)] \tag{2.60}
\]

**Boltzmann’s \( H \)-Theorem:** If \( f(p, t) \) satisfies the Boltzmann’s transport equation, then

\[
\frac{dH(t)}{dt} \leq 0. \tag{2.61}
\]
CHAPTER 2. KINETIC THEORY

Proof: By substituting (2.58) in (2.60) we obtain

\[
\frac{dH(t)}{dt} = \int dp \frac{\partial}{\partial t} [f(p, t) \log f(p, t)] = \int dp_1 dp_2 dp'_1 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} \times \\
\times (f(p'_1) f(p'_2) - f(p_1) f(p_2)) [1 + \log f(p_1)]. \tag{2.62}
\]

which is equivalent to

\[
\frac{dH(t)}{dt} = \int dp_1 dp_2 dp'_1 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} \times \\
\times (f(p'_1) f(p'_2) - f(p_1) f(p_2)) [1 + \log f(p_2)]. \tag{2.63}
\]

Note that (2.62) and (2.63) only differ in the argument of the logarithm. Thus we can add the two equation to obtain

\[
2 \frac{dH(t)}{dt} = \int dp_1 dp_2 dp'_1 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} \times \\
\times (f(p'_1) f(p'_2) - f(p_1) f(p_2)) [2 + \log (f(p_1) f(p_2))]. \tag{2.64}
\]

Moreover \(dp_1 dp_2 dp'_1 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2}\) is invariant under interchange of \(p_1, p_2\) and \(p'_1, p'_2\) because of the time invariance of the equations of motion, i.e.

\[
2 \frac{dH(t)}{dt} = \int dp_1 dp_2 dp'_1 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} \times \\
\times (f(p_1) f(p_2) - f(p'_1) f(p'_2)) [2 + \log (f(p'_1) f(p'_2))]. \tag{2.65}
\]

By adding together (2.64) and (2.65) we get

\[
\frac{dH(t)}{dt} = -\frac{1}{4} \int dp_1 dp_2 dp'_1 dp'_2 P_{p_1, p_2 \rightarrow p'_1, p'_2} \times \\
\times (f(p_1) f(p_2) - f(p'_1) f(p'_2)) [\log (f(p_1) f(p_2)) - \log (f(p'_1) f(p'_2))]. \tag{2.66}
\]

Since the log function is monotonically increasing we conclude that

\[
(f(p_1) f(p_2) - f(p'_1) f(p'_2)) [\log (f(p_1) f(p_2)) - \log (f(p'_1) f(p'_2))] \geq 0 \tag{2.67}
\]

as well as

\[
\frac{dH(t)}{dt} \leq 0. \tag{2.68}
\]

It should also be clear that \(\frac{dH(t)}{dt} = 0\) if an only if

\[
f_{eq}(p_1) f_{eq}(p_2) = f_{eq}(p'_1) f_{eq}(p'_2), \tag{2.69}
\]
which corresponds on an equilibrium distribution \( f_{eq}(p, t) \) yet to be calculated. Therefore,

\[
\frac{\partial f_{eq}(p, t)}{\partial t} = 0 \quad \Rightarrow \quad \frac{dH(t)}{dt} = \int d\mathbf{p} \frac{\partial f(p, t)}{\partial t} [1 + \log f(p, t)] = 0 \quad \Rightarrow \quad f_{eq}(\mathbf{p}_1) f_{eq}(\mathbf{p}_2) = f_{eq}(\mathbf{p}_1') f_{eq}(\mathbf{p}_2')
\] (2.70)

In addition it is easy to check that

\[
f_{eq}(\mathbf{p}_1, t) f_{eq}(\mathbf{p}_2, t) = f_{eq}(\mathbf{p}_1', t) f_{eq}(\mathbf{p}_2', t) \Rightarrow \frac{\partial f_{eq}(\mathbf{p}, t)}{\partial t} = \int d\mathbf{p}_2 d\mathbf{p}_1' d\mathbf{p}_2' \times
\]

\[
\times \mathcal{P}_{\mathbf{p}_1 \rightarrow \mathbf{p}'_1, \mathbf{p}'_2} ([f(\mathbf{p}_1') f(\mathbf{p}_2') - f(\mathbf{p}_1) f(\mathbf{p}_2)]) = 0
\] (2.71)

This proves that (2.69) and (2.58) are equivalent:

\[
\frac{\partial f}{\partial t} = 0 \Leftrightarrow f_{eq}(\mathbf{p}_1, t) f_{eq}(\mathbf{p}_2, t) = f_{eq}(\mathbf{p}_1', t) f_{eq}(\mathbf{p}_2', t).
\] (2.72)

### 2.5 Equilibrium distribution

To find equilibrium distribution (also called Maxwell-Boltzmann distribution) we take logarithm of both sides of (2.69):

\[
\log f_{eq}(\mathbf{p}_1) + \log f_{eq}(\mathbf{p}_2) = \log f_{eq}(\mathbf{p}_1') + \log f_{eq}(\mathbf{p}_2').
\] (2.73)

If \( \chi_i(p) \) are the conserved quantities such that

\[
\chi_i(\mathbf{p}_1) + \chi_i(\mathbf{p}_2) = \chi_i(\mathbf{p}_1') + \chi_i(\mathbf{p}_2')
\] (2.74)

during binary collisions, then the equilibrium solution

\[
\log f_{eq}(\mathbf{p}) = C + \chi_1(\mathbf{p}) + \chi_2(\mathbf{p}) + ...
\] (2.75)

would also respect the conservation laws. If the conserved quantities are three components of momentum and energy, then

\[
\log f_{eq}(\mathbf{p}) = B_1 + B_2 p_x + B_3 p_y + B_4 p_z + B_5 \mathbf{p}^2
\] (2.76)

or

\[
\log f_{eq}(\mathbf{p}) = -A(\mathbf{p} - \bar{\mathbf{p}})^2 + \log C
\] (2.77)
CHAPTER 2. KINETIC THEORY

where

\[ B_1 = \log C - A \left( \bar{p}_x^2 + \bar{p}_y^2 + \bar{p}_z^2 \right) \]  \hspace{1cm} (2.78)
\[ B_2 = 2Ap_x \]  \hspace{1cm} (2.79)
\[ B_3 = 2Ap_y \]  \hspace{1cm} (2.80)
\[ B_4 = 2Ap_z \]  \hspace{1cm} (2.81)
\[ B_5 = -A. \]  \hspace{1cm} (2.82)

Then the equilibrium distribution is

\[ f_{eq}(p) = Ce^{-A(p - \bar{p})^2}, \]  \hspace{1cm} (2.83)

where the number density of particles

\[ \frac{N}{V} = \int f_{eq}(p) dp = \int Ce^{-A(p - \bar{p})^2} dp = C \int e^{-Ap^2} dp = C \left( \frac{\pi}{A} \right)^{\frac{3}{2}} \]  \hspace{1cm} (2.84)

and the average momentum

\[ \langle p \rangle = \frac{\int pCe^{-A(p - \bar{p})^2} dp}{\int Ce^{-A(p - \bar{p})^2} dp} = \frac{\int (p + \bar{p}) Ce^{-Ap^2} dp}{\int Ce^{-Ap^2} dp} = \bar{p}. \]  \hspace{1cm} (2.85)

If there is not translational motion or in a reference frame moving along with fluid, i.e. \( \bar{p} = 0 \), the average energy

\[ \varepsilon = \left\langle \frac{p^2}{2m} \right\rangle = \frac{\int \frac{p^2}{2m} Ce^{-Ap^2} dp}{\int Ce^{-Ap^2} dp} = \frac{\int \frac{p^2}{2m} Ce^{-Ap^2} 4\pi p^2 dp}{C \left( \frac{\pi}{A} \right)^{\frac{3}{2}}} = \frac{3}{4Am}. \]  \hspace{1cm} (2.86)

From (2.84) and (2.86) we get expressions for

\[ C = \frac{N}{V} \left( \frac{3}{4\pi\varepsilon m} \right)^{3/2} \]  \hspace{1cm} (2.87)

and

\[ A = \frac{3}{4\varepsilon m}. \]  \hspace{1cm} (2.88)

in terms of physical quantities \( N, V \) and \( \varepsilon \).

To relate these quantities to directly measurable quantities we calculate the pressure, i.e. average force per unit area. Consider a force on wall in a \( y-z \) direction. When a particle hits the wall its momentum changes by \( 2p_x \), i.e. impulse per particle is \( 2p_x \). The number of particles with with velocity
\( \frac{p_x}{m} > 0 \) reflected by the wall per unit time \( \delta t \) per unit area \( \delta A \) is given by 
\( \delta A \delta t \frac{p_x}{m} f_{eq}(p) dp \). Therefore the pressures (assuming \( \bar{p} = 0 \)) is

\[
P = \frac{\text{Impulse}}{\text{Area} \times \text{Time}} = \frac{\int_{p_x > 0} 2p_x \delta A \delta t \frac{p_x}{m} f_{eq}(p) dp}{\delta A \delta t} = 4 \int_{p_x > 0} \frac{p_x^2}{2m} Ce^{-Ap^2} dp.
\]

By combining together with (2.86) we arrive to the equation of state

\[
P = \frac{2}{3} \frac{N \varepsilon}{V}.
\] (2.90)

Remember that the temperature is defined experimentally as

\[
T = \frac{PV}{Nk}
\] (2.91)

which implies

\[
\varepsilon = \frac{3}{2} kT.
\] (2.92)

Then the equilibrium distribution can be written as

\[
f_{eq}(p) = \frac{N}{V(2\pi mkT)^{3/2}} e^{-\left(\frac{p - \bar{p}}{2mkT}\right)^2}.
\] (2.93)

The thermodynamical quantities can be derived as follows:

- The internal energy is the overall energy of particles, i.e.

\[
U = N \varepsilon = \frac{3}{2} NkT.
\] (2.94)

- The First Law of thermodynamics defines heat absorbed by the system

\[
dQ = dU + PdV
\] (2.95)

where the work done by particles \( dW = PdV \) follows from the definition of pressure (2.89).
- The heat capacity at constant volume

\[ C_V = \left( \frac{\partial U}{\partial T} \right)_V = \frac{3}{2} Nk. \quad (2.96) \]

- The Second Law of thermodynamics is the Boltzmann’s \( H \)-theorem where the non-increasing quantity \( H \) defined a non-decreasing thermodynamical entropy, i.e.

\[ S = -HVk. \quad (2.97) \]

**Loschmidt’s Paradox:** \( H \)-THEOREM SINGLES OUT A PREFERRED DIRECTION OF TIME (ARROW OF TIME) AND THUS, MUST BE INCONSISTENT WITH TIME REVERSAL INVARIANCE OF FUNDAMENTAL LAWS. HOW CAN IT BE POSSIBLE?