PX449 Kinetic Theory

Warwick Physics Society

2024

Contents

0 Motivation and Introduction **5**
0.1 Notation 5
0.1.1 Vectors 5
0.1.2 Index notation 5
0.1.3 Quick review 6
0.2 Tensors 6
0.2.1 Vector and tensor operations 6

1 Equation of Continuity, Phase Space and Liouville's theorem **9**
1.1 Equation of Continuity 9
1.2 Phase Space 9
1.3 Liouville Equation and Theorem 10
1.4 BBGKY Hierarchy 12
1.5 Quantities in SPDFs 15

2 Collisions and Scattering **17**
2.1 Boltzmann equation for Impurity Scattering 17
2.1.1 Microcanonical Ensemble 21
2.1.2 Canonical Ensemble 22
2.2 Binary Collisions 22
2.2.1 Elastic binary collisions 23

3 Reversibility, H Theorem and Equilibrium **25**
3.1 Reversibility and the Boltzmann Equation 25
3.2 The H Theorem 26
3.3 Equilibrium and the Maxwell Distribution 28

4 Local Equilibrium and Transport **31**
4.1 Local Equilibrium 31
4.1.1 Advective term from f_0 33
4.1.2 Relaxation Time Approximation 34
4.2 Impurity Scattering and Diffusion 35
4.2.1 Second moments of the Maxwell Distribution 35
4.2.2 Einstein Relation 36
4.2.3 Quasi-Static Evolution 36

4.3 Stress and Viscosity . . . 37
  4.3.1 Quasi-Static Evolution . . . 39
4.4 Thermal Conductivity . . . 39
4.5 Shortcomings . . . 40

5 Diffusion **41**
5.1 Quantum Fermi Gas . . . 41
  5.1.1 Scattering and Equilibrium . . . 42
  5.1.2 Diffusion, Conductance and Einstein's Law . . . 44
5.2 Weak Scattering, Momentum diffusion and Fokker-Planck Equation . . . 46
  5.2.1 Review of the Diffusion Equation . . . 46
  5.2.2 Momentum Diffusion . . . 50
5.3 Black-Scholes Equation . . . 50

6 Kinetic Theory of Plasma **55**

7 Credits **57**

Chapter 0

Motivation and Introduction

This module aims to study a branch of physics known as *Kinetic Theory*, i.e. how to understand large systems of (mostly) indistinguishable particles. It is important to recall equations and ideas you may have encountered from other modules.

This document is a revision guide which aims to make the module clearer when revising. It closely follows the notation, content and order of the lecture notes for the year 2023/24 for this module. It also takes into account content mentioned in live lectures that isn't in the notes. This should hopefully make these notes a more comprehensive revision tool.

If there are any issues please email `physsocacademics@gmail.com`, with a screenshot of the issue and its location in the document.

0.1 Notation

0.1.1 Vectors

All vectors are denoted with an arrow → above the symbol like so: $\vec{p} = (p_1, p_2, \dots, p_n)$ means the vector quantity of momentum, and we define the corresponding no-arrowed, no sub/superscripted symbol to be its magnitude in the $\mathbb{R}^n$ norm, i.e.

$$p = |\vec{p}| = \left(\sum_{i=1}^{n} p_i^2 \right)^{1/2},$$

in contrast to other notes which may use a boldface.

0.1.2 Index notation

This module uses index notation, that is the stuff you might have seen in *PX3A3 Electrodynamics* and *PX3A5 The Standard Model* that looks like $x_\mu x^\mu$. In general, the position of the μ does matter and denotes the index of a 4-vector, but for this module, the position does not matter and we will generally be using subscripts, i.e. x_μ.

There is also some use of the Einstein Summation Convention, where $\sum$ are explicit running indices are all dropped. Therefore the dot product of 2 vectors $\vec{x}, \vec{p}$ can be written as $x_i p_i = \sum_{i=1}^{n} x_i p_i$.

0.1.3 Quick review

Equation of Continuity. ***PX285 Hamiltonian and Fluid Mechanics.*** Also known as the Continuity Equation, it describes *conservation of mass*, or of some flux. Consider some closed volume Ω immersed in a flux density $\vec{j}$. Let $J(t)$ denote the amount of stuff (flux) in Ω at time t. The amount of stuff exiting the boundary $(\partial\Omega)$, satisfies

$$\frac{\mathrm{d}J}{\mathrm{d}t} + \vec{\nabla} \cdot \vec{j} = 0. \tag{1}$$

This means that the quantity J is **conserved**.

The Hamiltonian. ***PX285 Hamiltonian and Fluid Mechanics.*** We denote the Hamiltonian as curly H: $\mathcal{H}$. In classical mechanics, this is defined as

$$\mathcal{H} = T(\vec{p}) + \Phi(\vec{r}) \tag{2}$$

where $\vec{p}$ is the momentum of one particle, and $\vec{r}$ is its position. T is the classical kinetic energy $p^2/2m$ and Φ is some potential. This is the **single-particle Hamiltonian** which is usually denoted as $\mathcal{H}_{(1)}(i)$ for the ith particle (just subscript every quantity with i). This extends to two particles etc.

0.2 Tensors

Definition 0.2.1. An **nth rank tensor** is a mathematical object with n independent indices.

For example X_μ is a first rank tensor because all its information can be accessed with 1 independent index μ. These are sometimes called vectors and denoted $\vec{X}$. Scalars are zero-rank tensors. $F^{\mu\nu}$ (the Faraday tensor from PX3A3) is a second-rank tensor.

0.2.1 Vector and tensor operations

This module is fairly maths heavy and involves a fair amount of switching between index notation and vector notation. The index notation should be easier to understand - it specifies the maths of one component, which you repeat for all other components. Vector notation can be however, complicated.

Definition 0.2.2. Let $\vec{u} \in \mathbb{R}^m, \vec{v} \in \mathbb{R}^n$. Then their **outer product** is defined by

$$\vec{u} \otimes \vec{v} = \mathbf{A} = \begin{bmatrix} u_1v_1 & u_1v_2 & \dots & u_1v_n \\ u_2v_1 & u_2v_2 & \dots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_mv_1 & u_mv_2 & \dots & u_mv_n \end{bmatrix} = A_{ij}$$

i.e. we have a binary operation $\otimes : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^{mn}$

Furthermore, the outer product is associative.

In this case, matrices and second-rank tensors are the same thing. These aren't strictly the same across all maths and physics (i.e., relativity), but for this module we will take it to be.

Another thing to ask if: let $u : \mathbb{R}^n \to \mathbb{R}$ be a scalar-valued function and $u : \mathbb{R}^n \to \mathbb{R}^n$ a vector-valued function. What is the difference between:

$$\vec{\nabla} u \text{ vs. } \vec{\nabla} \vec{u} \tag{3}$$

We define $(\vec{\nabla})^T = (\partial_1, \partial_2, \ldots, \partial_n)$. The second term is actually a **second-rank tensor** and the first expression is a **first-rank tensor**.

$$\vec{\nabla} u = \begin{bmatrix} \partial_1 u \\ \partial_2 u \\ \vdots \\ \partial_n u \end{bmatrix} \text{ vs. } \vec{\nabla}\vec{u} = \vec{\nabla} \otimes \vec{u} = \begin{bmatrix} \partial_1 u_1 & \partial_1 u_2 & \ldots & \partial_1 u_n \\ \partial_2 u_1 & \partial_2 u_2 & \ldots & \partial_2 u_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n u_1 & \partial_n u_2 & \ldots & \partial_n u_n \end{bmatrix} \tag{4}$$

Since we generally dealing with everything as vectors, we use the second form the most. It will be the case in this module that 'multiplication' of vectors (written as $\vec{x}\vec{x}$) is implied to be the outer product ($\vec{x}\vec{x} := \vec{x} \otimes \vec{x}$) and the dot and cross products will be written explicitly.

Chapter 1

Equation of Continuity, Phase Space and Liouville's theorem

1.1 Equation of Continuity

We know the regular continuity equation for a conserved quantity in (1). If we have a *source* or *sink* term, the right-hand-side may be non-zero, and equal to some function $Q(t, \vec{x})$ which is the amount of J generated per unit volume per unit time. Being more explicit:

$$\frac{\mathrm{d}J}{\mathrm{d}t} + \vec{\nabla} \cdot \vec{j} = Q(t, \vec{x}) \tag{1.1}$$

However, let us stick with the conserved version (1).

1.2 Phase Space

Consider a system of N interacting particles in 3 spatial dimensions with position and momenta $(\vec{r}_i, \vec{p}_i) = (r_{i,1}, r_{i,2}, r_{i,3}, p_{i,1}, p_{i,2}, p_{i,3})$.

Definition 1.2.1. Phase space is the $6N-$dimensional manifold which contains all possible configurations of position and momenta for a system.

This definition need not be memorised, nor is it exact[1]. It is also not necessary to know what a manifold is.

Definition 1.2.2. The evolution of a system can be described by a set of **phase-space trajectories**, which are lines through phase space.

The motivation to use these definitions is that we can recast all classical mechanics systems strictly in terms of position and momenta. As such, we may study them using Hamiltonian mechanics.

Furthermore, we wish to have a statistical description of our phase spaces, much like in previous modules covering statistical mechanics (*PX3A7 Statistical Physics, PX284 Statistical Mechanics, Electromagnetic Theory and Optics*) where we have worked with statistical distributions of particles. This motivates the following definition:

---

[1] In classical mechanics, it is a symplectic manifold

Definition 1.2.3. The **N-particle distribution function** $\mathcal{F}_N$ for our system of N particles is

$$\mathcal{F}_N(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_N, \vec{p}_N, t) \coloneqq \mathcal{F}_N(\vec{r}^N, \vec{p}^N, t) \tag{1.2}$$

This is a probability distribution function which gives the probability of the system to be in state $(\vec{r}_i, \vec{p}_i) = (r_{i,1}, r_{i,2}, r_{i,3}, p_{i,1}, p_{i,2}, p_{i,3})$ for every particle i.

If we have a probability distribution, it must satisfy a **normalisation condition**.

Lemma 1.2.1. *Normalisation condition for $\mathcal{F}_N$ is:*

$$\int\int \mathcal{F}_N(\vec{r}^N, \vec{p}^N, t) d\vec{r}^N d\vec{p}^N = 1 \tag{1.3}$$

where we integrate over all position and momentum degree of freedoms.

This integral is the equivalent of saying it is guaranteed for the system to reach every position and momentum it needs to (it might not reach every possible value of momentum and position however).

Most of the rest of this module depends on this function in some way.

1.3 Liouville Equation and Theorem

Equivalent definitions of Liouville's Theorem (the physics-y ones anyways)

- Hamiltonian dynamics conserves phase space volume.
- Phase space is incompressible under Hamiltonian flow.
- The phase space distribution function is constant along Hamiltonian trajectories

We will soon construct and prove Liouville's theorem.

As mentioned before, the definitions of phase space motivate the use of Hamilton's equations[2] which are

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial r_\mu}\bigg|_{p_\alpha} \qquad \frac{\mathrm{d}r_\mu}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_\mu}\bigg|_{r_\alpha} \tag{1.4}$$

where the Hamiltonian is a function of all particles, using the same notation for $\mathcal{F}_N$:

$$\mathcal{H} = \mathcal{H}(\vec{r}^N, \vec{p}^N, t) = T(\vec{p}^N) + \Phi(\vec{r}^N)$$

where again T is the total classical kinetic energy and Φ is the total potential energy *which depends only on particle positions.* We will shortly return to the usefulness of the Hamiltonian in this context.

Definition 1.3.1. A **phase-space position** is an element of phase space. This is denoted by the 2-tuple $(\vec{r}, \vec{p})$ (or $(\vec{r}^N, \vec{p}^N)$ for our N particle system).

---

[2] PX285 Hamiltonian and Fluid Mechanics

We will refer to this as simply 'position' when it is clear from context.

Definition 1.3.2. A **phase-space velocity** is the total time derivative of position in phase space. This is denoted by the 2-tuple $(\dot{\vec{r}}, \dot{\vec{p}})$ (or $(\dot{\vec{r}}^N, \dot{\vec{p}}^N)$ for our N particle system).

We will refer to this as simply 'velocity' when it is clear from context.

Now if we have phase-space velocity and we have trajectories, each position is like a position vector. The velocity vectors are always tangential to a phase-space trajectory at a fixed time t. This has the analogue to a fluid: fluid velocity vectors are tangential to streamlines, and streamlines describe the trajectory of a fluid particle (in the Eulerian picture). What we have is effectively a **fluid**!

Lemma 1.3.1. *The fluid (aka Hamiltonian) flow in phase space is* ***divergence-free or incompressible****.*

Proof. Remember that a fluid flow is incompressible if the divergence of the velocity, $\vec{\nabla} \cdot \vec{u} = 0$. Here, our velocity is the phase-space velocity. We substitute for phase-space velocity directly:

$$\vec{\nabla} \cdot \left(\dot{\vec{r}}^N, \dot{\vec{p}}^N\right) = \frac{\partial \dot{r}_\mu}{\partial r_\mu} + \frac{\partial \dot{p}_\mu}{\partial p_\mu} = \frac{\partial^2 \mathcal{H}}{\partial r_\mu \, \partial p_\mu} - \frac{\partial^2 \mathcal{H}}{\partial p_\mu \, \partial r_\mu} = 0$$

where in the last step we use commutativity of second-order partial derivatives to instantly equate it to 0. □

Remark. *Note the summation convention. We do this over each component of position and velocity.*

Remark. *Observe this is one of the formulations of Liouville's theorem.*

We can actually go and derive Liouville's equation with this, which is a key result to remember

Theorem 1.3.2. *Liouville's Equation*

$$\frac{\partial \mathcal{F}_N}{\partial t} + \dot{r}_\mu \frac{\partial \mathcal{F}_N}{\partial r_\mu} + \dot{p}_\mu \frac{\partial \mathcal{F}_N}{\partial p_\mu} = 0 \tag{1.5}$$

Proof. The probability is conserved and so obeys a continuity equation

$$\frac{\partial \mathcal{F}_N}{\partial t} + \vec{\nabla} \cdot \left[\mathcal{F}_N \left(\dot{r}_\mu, \dot{p}_\mu\right)\right] = 0$$

Using the product rule on the derivative and using Lemma 1.3.1 we arrive at

$$\frac{\partial \mathcal{F}_N}{\partial t} + \left(\dot{r}_\mu, \dot{p}_\mu\right) \cdot \vec{\nabla} \mathcal{F}_N = 0$$

Note that

$$\vec{\nabla} \mathcal{F}_N = \left(\frac{\partial \mathcal{F}_N}{\partial r_\mu}, \frac{\partial \mathcal{F}_N}{\partial p_\mu}\right) \tag{1.6}$$

So expanding the dot product above, we have

$$\frac{\partial \mathcal{F}_N}{\partial t} + \left(\dot{r}_\mu, \dot{p}_\mu\right) \cdot \left(\frac{\partial \mathcal{F}_N}{\partial r_\mu}, \frac{\partial \mathcal{F}_N}{\partial p_\mu}\right) = 0$$

which directly gives us Liouville's equation □

This leads us directly to

Theorem 1.3.3. *Liouville's Theorem: the probability following a phase-space trajectory is conserved, i.e.*

$$\mathcal{F}_N(\vec{r}^N, \vec{p}^N, t) = \mathcal{F}_N(\vec{r}_0^N, \vec{p}_0^N, 0) \tag{1.7}$$

where the RHS is the probability density at the initial time $t = 0$.

Proof. The previous Theorem 1.3.2 tells us $\mathcal{F}_N(\vec{r}^N, \vec{p}^N, t)$is conserved. To ensure a well-posed problem to a classical mechanics problem (i.e. Hamilton's equations 1.4, we need *initial conditions*) which will be specified in terms of momenta and positions. This means the probability density must be the same as it was specified as the start of the system evolution.

(NON-EXAMINABLE) Mathematically, the differential phase space volume occupied by two different elements in phase space are the same. See this LibreText for more information.

□

1.4 BBGKY Hierarchy

Finding the N-particle density is hard. We want to find the single-particle density, the probability density that ANY particle has specified position and momentum, independent of the others (this is kind of like always needing to find the partition function for statistical mechanics).

Definition 1.4.1. The **single-particle density function** (SPDF) $f(\vec{r}, \vec{p}, t)$ is defined by the singular contributions:

$$f(\vec{r}, \vec{p}, t) = \int \mathcal{F}_N\left(\vec{r}^N, \vec{p}^N, t\right) \sum_{i=1}^{N} \left(\delta\left(\vec{r}_i - \vec{r}\right) \delta\left(\vec{p}_i - \vec{p}\right)\right) d\vec{r}^N d\vec{p}^N \tag{1.8}$$

The δ functions pick out singular contributions from each particle i.

For **indistinguishable** particles, we can say that the single particle density is N times the probability density for particle 1[3]

Then f reduces to:

$$f(\vec{r}, \vec{p}, t) = N \int \mathcal{F}_N\left(\vec{r}, \vec{p}, \vec{r}_2, \vec{p}_2 \dots \vec{r}_N, \vec{p}_N, t\right) d\vec{r}_2 d\vec{p}_2 \dots d\vec{r}_N d\vec{p}_N \tag{1.9}$$

Now, since f is a probability density, it must obey some normalisation condition, specifically:

$$\int f(\vec{r}, \vec{p}, t) d\vec{p} d\vec{r} = N \tag{1.10}$$

Lemma 1.4.1.

$$\int f(\vec{r}, \vec{p}, t) d\vec{p} = n(\vec{r}) \tag{1.11}$$

where $n(\vec{r})$ is the spatial number density of particles.

---

[3] They are indistinguishable, so labelling a particle doesn't matter, but it simply means we can choose just any one of them for consideration

Now we have f, we want to find its time-evolution like we did for $\mathcal{F}_N$. To do this we make a series of approximations:

- Ignore any 3+ body interactions. Only consider single-particle and pairwise-particle interactions.
- Treat the 2-body Hamiltonian as an additional perturbation in the potential of the single-particle Hamiltonian

Let us formalise this. Write

$$\mathcal{H}(\vec{r}^N, \vec{p}^N) = \sum_i \mathcal{H}_{(1)}(i) + \sum_{j>i} \mathcal{H}_{(2)}(i,j) \tag{1.12}$$

where $\mathcal{H}_{(1)}(i) = \mathcal{H}_{(1)}(\vec{r}_i, \vec{p}_i)$ and $\mathcal{H}_{(2)}(i,j) = \mathcal{H}_{(2)}(j,i) = \mathcal{H}_{(2)}(\vec{r}_i, \vec{p}_i, \vec{r}_j, \vec{p}_j)$ and $j > i$ to avoid double counting. Additionally $\mathcal{H}_{(1)}(i) = T(\vec{p}_i) + U_{\text{ext}}(\vec{r}_i)$ and $\mathcal{H}_{(2)} = V(\vec{r}_i - \vec{r}_j)$.

We will now substitute this into Liouville's equation (1.3.2).

Theorem 1.4.2. *$f(\vec{r}, \vec{p})$ obeys a single-particle Liouville equation, i.e. it satisfies Theorem 1.3.2.*

Remark. *THIS WAS A QUESTION IN THE PX449 2023 EXAM!*

Before we prove the theorem though, we must deal with being able to partially-differentiate Eq.(1.9). We state without proof:

Lemma 1.4.3. *Leibniz integral rule*

$$\frac{\partial}{\partial z}\int_{a(z)}^{b(z)} f(x,z)dx = \int_{a(z)}^{b(z)} \frac{\partial f}{\partial z}dx + f(b(z), z)\frac{\partial b}{\partial z} - f(a(z), z)\frac{\partial a}{\partial z}$$

where the function f here is unrelated to $f(\vec{r}, \vec{p}, t)$

Then we replace $z \to t$ and ignore the functions $a(z \to t), b(z \to t)$ since the original integral in Eq. (1.9) is over all position and momenta:

$$\frac{\partial f(\vec{r}, \vec{p}, t)}{\partial t} = N \int \frac{\partial \mathcal{F}_N}{\partial t} d\vec{r}_2 d\vec{p}_2 \dots d\vec{r}_N d\vec{p}_N \tag{1.13}$$

where the derivative terms are 0 since we assume that position and momenta are time independent - they are independent degrees of freedom.

Now we return to the proof of Theorem 1.4.2

Proof. First, we rewrite Liouville's equation in terms of the Hamiltonian using Hamilton's equations and remove summation convention:

$$\frac{\partial \mathcal{F}_N}{\partial t} + \sum_i \left(\frac{\partial}{\partial \vec{r}_i} \cdot \left(\frac{\partial \mathcal{H}}{\partial \vec{p}_i} \mathcal{F}_N \right) - \frac{\partial}{\partial \vec{p}_i} \cdot \left(\frac{\partial \mathcal{H}}{\partial \vec{r}_i} \mathcal{F}_N \right) \right) = 0 \tag{1.14}$$

We will substitute Eq. (1.12) into Eq. (1.14).

$$\begin{aligned}\frac{\partial \mathcal{F}_N}{\partial t} &+ \sum_i \left(\frac{\partial}{\partial \vec{r}_i} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(i)}{\partial \vec{p}_i} \mathcal{F}_N \right) - \frac{\partial}{\partial \vec{p}_i} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(i)}{\partial \vec{r}_i} \mathcal{F}_N \right) \right) \\ &+ \sum_{i<j} \left(\frac{\partial}{\partial \vec{r}_i} \cdot \left(\frac{\partial \mathcal{H}_{(2)}(ij)}{\partial \vec{p}_i} \mathcal{F}_N \right) - \frac{\partial}{\partial \vec{p}_i} \cdot \left(\frac{\partial \mathcal{H}_{(2)}(ij)}{\partial \vec{r}_i} \mathcal{F}_N \right) \right) = 0\end{aligned}$$

Observe we can find an expression for simply $\frac{\partial \mathcal{F}_N}{\partial t}$ by rearranging the two series terms to the other side, and so substitute directly into Eq. (1.13)

$$\begin{aligned}\frac{\partial f(\vec{r},\vec{p},t)}{\partial t} = -N \int \sum_i & \left(\frac{\partial}{\partial \vec{r}_i} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(i)}{\partial \vec{p}_i} \mathcal{F}_N \right) - \frac{\partial}{\partial \vec{p}_i} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(i)}{\partial \vec{r}_i} \mathcal{F}_N \right) \right) \\ &+ \sum_{i<j} \left(\frac{\partial}{\partial \vec{r}_i} \cdot \left(\frac{\partial \mathcal{H}_{(2)}(ij)}{\partial \vec{p}_i} \mathcal{F}_N \right) - \frac{\partial}{\partial \vec{p}_i} \cdot \left(\frac{\partial \mathcal{H}_{(2)}(ij)}{\partial \vec{r}_i} \mathcal{F}_N \right) \right) d\vec{p}_2 \dots d\vec{r}_N d\vec{p}_N\end{aligned}$$

Integrating out all the $i \geq 2$ terms we see they are 0. This leaves the following expression:

$$\begin{aligned}\frac{\partial f(\vec{r},\vec{p})}{\partial t} + \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{p}} f(\vec{r},\vec{p}) \right) - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{r}} f(\vec{r},\vec{p}) \right) \\ + \int d\vec{r}' d\vec{p}' \left\{ \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(2)}\left(\vec{r},\vec{p},\vec{r}',\vec{p}'\right)}{\partial \vec{p}} f_{(2)}\left(\vec{r},\vec{p},\vec{r}',\vec{p}'\right) \right) \right. \\ \left. - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(2)}\left(\vec{r},\vec{p},\vec{r}',\vec{p}'\right)}{\partial \vec{r}} f_{(2)}\left(\vec{r},\vec{p},\vec{r}',\vec{p}'\right) \right) \right\} = 0.\end{aligned}$$

Ignoring the $\mathcal{H}_{(2)}$ terms. we have

$$\boxed{\frac{\partial f(\vec{r},\vec{p})}{\partial t} + \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{p}} f(\vec{r},\vec{p}) \right) - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{r}} f(\vec{r},\vec{p}) \right) = 0 \qquad (1.15)}$$

This is exactly Liouville's equation. □

Remark. *The above expression in terms of $f(\vec{r},\vec{p})$ and $\mathcal{H}_{(1)}$ is also known as the **Collisionless Boltzmann Equation** since it ignores all interactions.*

Observe that the $\mathcal{H}_{(2)}$ terms have a pairwise particle density $f_{(2)}$, but that is involved in the time derivative of $\frac{\partial f}{\partial t}$. Repeating this, we can find $\frac{\partial f_{(2)}}{\partial t}$ in terms of third-particle densities and Hamiltonians etc, all the way until we go back to $\mathcal{F}_N$.

This is the BBGKY hierachy

Escaping from the hierarchy involves approximating some $f_{(k+1)}$ as a combination of lower order densities, but this must be done carefully to avoid physics-breaking shenanigans.

1.5 Quantities in SPDFs

SPDF is useful because we can write physical quantities in terms of it. It is much like how finding the partition function allows us to determine any thermodynamic quantity like energy and heat capacity. $f(\vec{r}, \vec{p})$ can do the same.

Recall these definitions for the exam

Definition 1.5.1. The spatial particle density

$$n(\vec{r}, t) = \int f(\vec{r}, \vec{p}, t) d\vec{p}$$

Definition 1.5.2. The first moment gives the particle flux density

$$j_\mu(\vec{r}, t) = \int \frac{p_\mu}{m} f(\vec{r}, \vec{p}, t) d\vec{p}$$

The mass flux density = momentum density in space.

Definition 1.5.3. The **Stress Tensor** is (minus) the flux density of momentum. It comes from a second moment.

$$\sigma_{\mu\nu}(\vec{r}) = -\int \frac{p_\mu p_\nu}{m} f d\vec{p}$$

Definition 1.5.4. The heat flux [density] has components given by a third moment.

$$Q_\mu(\vec{r}) = \int \frac{p_\mu}{m} \frac{p^2}{2m} f d\vec{p}$$

Chapter 2

Collisions and Scattering

Unless otherwise indicated, we use the single-particle Hamiltonian for Newtonian particles, defined by

Definition 2.0.1. The **single-particle Hamiltonian** $\mathcal{H}_{(1)}$ is defined as

$$\mathcal{H}_{(1)} = \frac{|\vec{p}|^2}{2m} + \Phi_{\text{ext}}(\vec{r})$$

We will assume that the potential Φ_{ext} is an external perturbation up ignoring the leading order terms.

The notation will however, remain in its most general form until simplification can be done using this specific version of $\mathcal{H}_{(1)}$.

2.1 Boltzmann equation for Impurity Scattering

An ensemble of colliding particles in 2D has a mean free path.

Definition 2.1.1. The **mean-free path** of a system of colliding particles in 2D (there may be multiple types of particles, but each particle of the same type is indistinguishable to the others of the same type) is defined by

$$\lambda = \frac{1}{\pi a^2 n_I} \tag{2.1}$$

where

- a is the **collision diameter** defined by the sum of radii of a particle and the particle it collides with (the 'scatterer')
- n_I is the number density of scatterer particles.

It may be that the scatterer particles are the same type as the colliding particles, in which case n_I is the number density of the system.

We can rewrite λ as

$$\lambda \simeq \frac{a}{\psi} \implies \psi \simeq n_I a^3 \ll 1 \tag{2.2}$$

Definition 2.1.2. ψ is the **volume fraction of scatterers** whose number density is n_I.

If we have a volume fraction, we can find a typical nearest neighbour distance between scattering events.

Definition 2.1.3. The **nearest-neighbour distance** ξ is the minimal distance of particles over the whole system between scattering events. This means

$$n\xi^3 \simeq 1 \implies \xi \simeq a\psi^{-1/3}$$

Then the ratio of the mean free path to nearest neighbour distance is

$$\frac{\lambda}{\xi} \simeq \psi^{-2/3} \tag{2.3}$$

Remark. *Useful to recall this ratio.*

Suppose $\lambda/a \gg 1$. This means the mean free path is much larger than the collision radius, so we are to treat our ensemble as colliding *points*.

Hence, in every collision, the energies of the colliding particles may change. This can result in sudden changes in direction of the phase-space trajectories of particles. In essence, the change in direction (acceleration) is driven by some change in energy. This is equivalent to adding a source or sink term in the Liouville Eq. (1.15). This form of the equation ignores the $\mathcal{H}_{(2)}$ terms, but the energy change does depend on them. We make the following definition:

Definition 2.1.4. The inner and outer scattering rates $S_{\text{in}}, S_{\text{out}}$ are the respective scattering rates (per unit volume in single particle phase space) in and out from phase-space position $(\vec{r}, \vec{p})$. We also call these the **gain** and **loss** functions respectively.

Then Eq. (1.15) becomes

$$\frac{\partial f(\vec{r},\vec{p})}{\partial t} + \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{p}} f(\vec{r},\vec{p}) \right) - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{r}} f(\vec{r},\vec{p}) \right) = S_{\text{in}}(\vec{r},\vec{p}) - S_{\text{out}}(\vec{r},\vec{p}) \tag{2.4}$$

We now switch terminology a bit and call our scatterer particles, *impurities*.

$S_{\text{out}}(\vec{r},\vec{p}) = n_I k_{\text{out}}$ where k_{out} is the scattering from a single impurity, i.e. the number of particles scattered due to one impurity in a phase space volume.

Now, imagine an impurity situated at $\vec{r}$, and an incoming circular beam of particles with momentum $\vec{p}$ directly heading towards the impurity. The beam has a density $f(\vec{r},\vec{p})$ and sweep into contact with the impurity at a rate k_{out}. We need a formula for k_{out}.

The amount of scattering depends on the number density and the incoming beam of particles. Consider some ring at the external of the beam $d\sigma$. The particles in this region scatter by different amounts depending on the radial distance from the centre of the beam, where inner particles scatter more than outer particles in a time dt.

Imagining our particles as hard spheres, we must project our area in the direction of scattering, i.e. in direction of $\vec{p}$. Namely, this motivates the following definition:

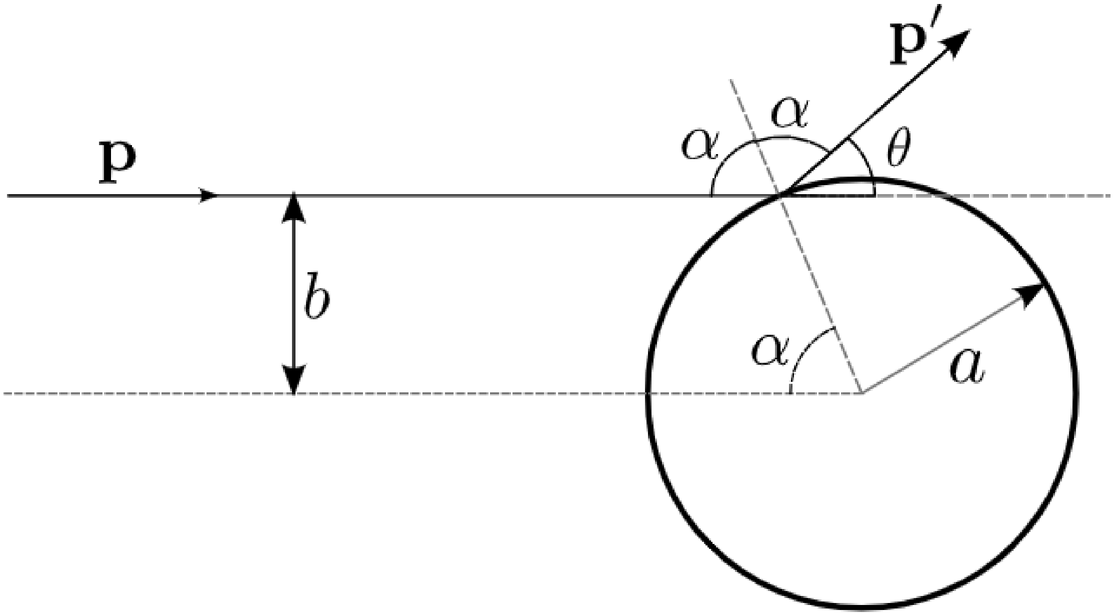

Figure 2.1: Plane view of impurity scattering

Definition 2.1.5. The **total scattering cross-section** σ is the measure that an interaction will occur.

It kind of behaves like a probability, however $\sigma > 1$ is very much allowed. A larger cross-section basically means a higher chance of an interaction happening.

cross-section is *not the same* as probability, but they are linked.

In our impurity scattering, the scattering that occurs depends on the collision diameter a where a larger a (e.g., larger impurities and/or larger particles) increase the likelihood of scattering. In particular, scattering can happen anywhere in a collision area, therefore $\sigma = \pi a^2$.

Therefore the rate is

$$k_{\text{out}} = \frac{|\vec{p}|}{m} f(\vec{r}, \vec{p}) \sigma$$

Definition 2.1.6. The **impact parameter** b is the centre-to-centre distance between the particle and impurity before scattering.

There is an observed distribution of outgoing directions and correspondingly a distribution of the cross-section over solid angles, characterised by the **differential scattering cross-section**.

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$$

We note that the ring $d\sigma = 2\pi b db$ and

Lemma 2.1.1. *$b = a\cos(\theta/2)$ for hard impurities.*

Proof. Consider the angle bisector from the centre of the impurity to the point of contact. See Fig. 2.1. The angle $180 - \theta = 2\alpha$. Using b to form a right-handed triangle, we observe $a \sin\alpha = b$. However then

$$\alpha = \frac{1}{2}(\pi - \theta) \implies \sin(\alpha) = \cos(\pm\theta/2) \implies b = a\cos(\theta/2)$$

□

Then $db = -a/2\sin(\theta/2)d\theta$ and $d\Omega = 2\pi\sin\theta d\theta$ (we are using solid angles as mentioned before). Notice

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left|\frac{2\pi b db}{2\pi\sin\theta d\theta}\right| = \frac{a^2}{4} \tag{2.5}$$

Remark. *The RHS has turned out to be independent of angle, so the scattered particles are* ***uniformly distributed****. Furthermore there is abuse of notation as this value is really the absolute value of the differential cross-section.*

Now, we can substitute back into k_{out} to obtain

$$S_{\text{out}} = n_I \frac{|\vec{p}|}{m} f(\vec{r}, \vec{p})\sigma = \int n_I \frac{|\vec{p}|}{m} f(\vec{r}, \vec{p}) \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} d\Omega \tag{2.6}$$

We want to write this in terms of a function $W(\vec{p}, \vec{p'})$ which incorporates **conservation of kinetic energy** (for elastic collisions) and is symmetric so $W(\vec{p}, \vec{p'}) = W(\vec{p}, \vec{p'})$, i.e. symmetric under exchange of $\vec{p} \leftrightarrow \vec{p'}$. This function W reads as the scattering that reached momentum $\vec{p'}$ from input momentum $\vec{p}$. Why do we want this? This allows us to find $S_{\text{out}}, S_{\text{in}}$ to generalise to any impurity scattering. To be more explicit, we formalise this by saying

$$S_{\text{out}} = \int W(\vec{p}, \vec{p'}) f(\vec{p}) d\vec{p'} \qquad\qquad S_{\text{in}} = \int W(\vec{p'}, \vec{p}) f(\vec{p'}) d\vec{p'} \tag{2.7}$$

Lemma 2.1.2. *The function W satisfying Eq. (2.7) is*

$$W(\vec{p}, \vec{p'}) = \frac{n_I}{m}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\delta\left(\frac{p^2}{2} - \frac{(p')^2}{2}\right)$$

Proof. Exchange $\vec{p} \leftrightarrow \vec{p'}$ and observe the δ-function is unchanged. Additionally, the δ-function selects the values of momenta corresponding to elastic collisions, where $p' = p$. Notice substituting this into Eqs.(2.7) recovers Eq. (2.6). □

Remark. *Using properties of δ functions, we can see:*

$$\delta\left(\frac{p^2}{2} - \frac{(p')^2}{2}\right) = \frac{1}{p'}\delta(p' - p) = \frac{p}{p'^2}\delta(p' - p) \tag{2.8}$$

But why did we do all of this? Refer back to the modified Liouville Eq. (2.4). We can now substitute Eqs. (2.7). We get

Boltzmann Equation

$$\frac{\partial f(\vec{r},\vec{p})}{\partial t}+\frac{\partial}{\partial \vec{r}}\cdot\left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{p}}f(\vec{r},\vec{p})\right)-\frac{\partial}{\partial \vec{p}}\cdot\left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{r}}f(\vec{r},\vec{p})\right)=C[f] \tag{2.9}$$

and we define

Definition 2.1.7. $C[f]$ as the **Collision Integral** which for elastic scattering is a linear functional, but generally is non-linear. In particular,

$$C[f]\propto\int\left(W\left(\vec{p'},\vec{p}\right)f\left(\vec{p'}\right)-W\left(\vec{p},\vec{p'}\right)f(\vec{p})\right)d\vec{p'}, \tag{2.10}$$

Lemma 2.1.3. *For hard-sphere impurity scattering,*

$$C[f]=\frac{\overline{f(p)}-f(\vec{p})}{\tau}$$

where τ is the average time between collisions, also known as the ***collision time****.*

Proof. The scattering in has come from other momenta. Since it is isotropic (W does not depend on direction of $\vec{p'}$, so only angular average can contribute) then we have a term proportional to $\overline{f(p)}$. Then the particles that come out depend on the probability distribution at momenta $\vec{p}$ so we have a term proportional to $f(\vec{p})$. Since impurity scattering only changes particle directions, it cannot change the angular average distribution so we must divide each by τ.

This is the reasoning from lectures. Exact derivation non-examinable. □

Comparing this $C[f]$ with Eq.(2.6), we see $-S_{\text{out}}=-f(\vec{p})/\tau$ which leads to the relation

$$\frac{1}{\tau}=n_I\sigma\frac{p}{m} \tag{2.11}$$

by comparing coefficients.

2.1.1 Microcanonical Ensemble

Kinetic energy conserved in elastic collisions, so the distribution of particle energy is fairly limited. We have a separate population for each possible value of $p=|\vec{p}|$ (and equivalently energy) and we can study each of them separately: the population at $p=p_0$ is characterised by a parameter p_0, or equivalently the kinetic energy $E=p_0^2/2m$, NOT temperature!

The equilibrium distribution is

$$f_{\text{MCE}}=f_0(p_0,n;\vec{r},\vec{p})=n\frac{1}{4\pi p^2}\delta(p-p_0) \tag{2.12}$$

where n is the density of this population.

2.1.2 Canonical Ensemble

Suppose we relax the idea of conservation of kinetic energy, e.g. through binary collisions as will be seen in Section 2.2. We want to keep using a prefactor of $1/\tau$ (makes life easier, we students already have enough), so we must average Eq.(2.11) with respect to speed $c = p/m$ so

$$\frac{1}{\tau} = n_I \sigma \bar{c} \tag{2.13}$$

where explicitly, $\bar{c} = \langle p/m \rangle$.

Remark. *For an ideal gas,*

$$\bar{c} = \sqrt{\frac{8k_BT}{m\pi}}$$

This may look familiar from PX154 Physics Foundations. *This is exactly the mean speed of an ideal gas following the Maxwell distribution. In fact, the Maxwell distribution*

$$P(c) = \left(\frac{m}{2\pi k_BT}\right)^{3/2} 4\pi c^2 e^{-mc^2/2k_BT}$$

is a valid equilibrium single particle density function, and is what non-equilibrium ideal particle distributions tend towards to. See Section 3.3 for more details.

2.2 Binary Collisions

We will define some notation for this section, and it may be different from the lecture notes. We consider two incident particles with momenta $\vec{p_1}, \vec{p_2}$, which collide and scatter with outgoing momenta $\vec{p_1}', \vec{p_2}'$ respectively. We reserve the symbols $\vec{p}, \vec{r}$ for the particles described by the distribution function $f(\vec{r}, \vec{p})$.

- We ignore collisions between three or more particles yet again. This is valid in cases like **dilute** or **rarefied** gases where separations between scatterings is $\gg a$.
- This ratio grows as $\psi^{-1/3}$ as in Definition 2.1.3.

Then Eq. (2.10) becomes

$$C[f] = \int W\left(\vec{p_1}' \to \vec{p_1}, \vec{p_2}' \to \vec{p_2}\right) f_{(2)}\left(\vec{p_1}', \vec{p_2}'\right) \tag{2.14}$$

$$-W\left(\vec{p_1} \to \vec{p_1}', \vec{p_2} \to \vec{p_2}'\right) f_{(2)}\left(\vec{p_1}, \vec{p_2}\right) d\vec{p_1} d\vec{p_2}' d\vec{p_1}' \tag{2.15}$$

To align it more consistently with notation used in lectures, this can be rewritten as

$$C[f] = \int W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) f_{(2)}\left(\vec{p_1}', \vec{p_2}'\right) \tag{2.16}$$

$$-W\left(\vec{p_1}\vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) f_{(2)}\left(\vec{p_1}, \vec{p_2}\right) d\vec{p_1} d\vec{p_2}' d\vec{p_1}'$$

To proceed, we use the following assumption:

Definition 2.2.1. The **assumption of Molecular Chaos** is that the momenta of colliding particles are uncorrelated, and independent of position. We set the scale of their independence to be a mean free path away. Then the probability of a collision can be written as a product of probabilities at each momenta, namely

$$f_{(2)}(\vec{p_1}, \vec{p_2}) \approx f\left(\vec{p_1}\right) f\left(\vec{p_2}\right) \tag{2.17}$$

Substitution into Eq. (2.16) for both the primed and unprimed momenta gives the

Boltzmann Collision Integral

$$\begin{aligned} C[f] = \int W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) f\left(\vec{p_1}'\right) f\left(\vec{p_2}'\right) \\ -W\left(\vec{p_1}\vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) f\left(\vec{p_1}\right) f\left(\vec{p_2}\right) d\vec{p_1} d\vec{p_2}' d\vec{p_1}' \end{aligned} \tag{2.18}$$

Substituting this $C[f]$ into Eq. (2.9) gives

Boltzmann Kinetic Equation

$$\begin{aligned} \frac{\partial f(\vec{r}, \vec{p})}{\partial t} + \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r}, \vec{p})}{\partial \vec{p}} f(\vec{r}, \vec{p}) \right) - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r}, \vec{p})}{\partial \vec{r}} f(\vec{r}, \vec{p}) \right) \\ = \int W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) f\left(\vec{p_1}'\right) f\left(\vec{p_2}'\right) \\ -W\left(\vec{p_1}\vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) f\left(\vec{p_1}\right) f\left(\vec{p_2}\right) d\vec{p_1} d\vec{p_2}' d\vec{p_1}' \end{aligned} \tag{2.19}$$

2.2.1 Elastic binary collisions

The scattering rate function W is constrained by conservation of energy and momentum. For particles of equal mass this is

Lemma 2.2.1. *Scattering rate function for equal mass particles is*

$$W\left(\vec{p}, \vec{p_1} \to \vec{p}', \vec{p_1}'\right) = A\delta\left(\vec{p} + \vec{p_1} - \vec{p}' - \vec{p_1}'\right) \delta\left(\frac{\vec{p}^2 + \vec{p}_1^2}{2} - \frac{\left(\vec{p}'\right)^2 + \left(\vec{p_1}'\right)^2}{2} \right) \tag{2.20}$$

where

$$A = \frac{1}{\mu^2} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} v_{relative}$$

*where $\mu = m/2$ is the **reduced mass of the collision pair**.*

Proof. NON-EXAMINABLE (PROBABLY) □

As in the impurity case, for hard particles this turns out to be the full expression for W.

Chapter 3

Reversibility, H Theorem and Equilibrium

3.1 Reversibility and the Boltzmann Equation

Definition 3.1.1. A **parity transformation** P is the inversion of all spatial directions, namely $\vec{p} \to -\vec{p}$ etc.

Definition 3.1.2. A **time-reversal transformation** T is the inversion of the direction of time, namely $\vec{p} \to \vec{p'}$ becomes $-\vec{p'} \to -\vec{p}$ etc.

Theorem 3.1.1. *The scattering rate function W is invariant under PT transformations , i.e.*

$$W\left(\vec{p_1}, \vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) = W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) \tag{3.1}$$

Proof. Let us apply T to W:

$$T : W\left(\vec{p_1}, \vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) \to W\left(-\vec{p_1}', -\vec{p_2}' \to -\vec{p_1}, -\vec{p_2}\right)$$

Then apply P transformation:

$$PT : W\left(\vec{p_1}, \vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) \to W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right)$$

This concludes the proof. □

This allows us to swap the primed and unprimed coordinates as we wish. If we assume the scattering is **reversible**, we can simplify Eq. (2.19) which we restate here for convenience.

$$\begin{aligned}\frac{\partial f(\vec{r},\vec{p})}{\partial t} + \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{p}} f(\vec{r},\vec{p})\right) - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}(\vec{r},\vec{p})}{\partial \vec{r}} f(\vec{r},\vec{p})\right) \\ = \int W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) f\left(\vec{p_1}'\right) f\left(\vec{p_2}'\right) \\ -W\left(\vec{p_1}\vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) f\left(\vec{p_1}\right) f\left(\vec{p_2}\right) d\vec{p_1} d\vec{p_2}' d\vec{p_1}'\end{aligned}$$

We can expand the partial derivatives using properties of dot products and using Lemma 1.3.1 that Hamiltonian flow is incompressible, this sets all divergence terms to 0, as they all involve

$\mathcal{H}_{(1)}$. Then on the RHS we can factorise out a W since by Theorem 3.1.1, they are the same W we obtain

$$\frac{\partial f}{\partial t} + \frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} \cdot \frac{\partial f}{\partial \vec{r}} - \frac{\partial \mathcal{H}_{(1)}}{\partial \vec{r}} \cdot \frac{\partial f}{\partial \vec{p}} = C[f] \tag{3.2}$$
$$= \int \left(f\left(\vec{p_1}'\right) f\left(\vec{p_2'}\right) - f(\vec{p_1}) f\left(\vec{p_2}\right)\right) W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) d\vec{p_2} d\vec{p_1}' d\vec{p_2'}$$

3.2 The H Theorem

We want to show the SPDF f always tends towards equilibrium.

Definition 3.2.1. The **Boltzmann-Gibbs/Shannon Entropy** is defined by

$$S = -k_B \sum_i p_i \ln p_i \tag{3.3}$$

where p_i is the probability of the ith microstate and the sum is taken over all microstates. This should be familiar from *PX285* and/or *PX3A7*.

Instead of discrete probabilities, we have a continuous SPDF over phase space, therefore

$$S(t) = -k_B \int f(\vec{r}, \vec{p}, t) \ln f(\vec{r}, \vec{p}, t) d\vec{r} d\vec{p} \tag{3.4}$$

This is usually written as a dimensionless entropy $H = -S(t)/k_B$ so

$$H(t) = \int f(\vec{r}, \vec{p}, t) \ln f(\vec{r}, \vec{p}, t) d\vec{r} d\vec{p} \tag{3.5}$$

Using Leibniz rule, and being implicit on variables,

$$\begin{aligned} \frac{dH}{dt} = \int \frac{\partial}{\partial t}(f \ln f) d\vec{r} d\vec{p} &= \int \left(\frac{\partial f}{\partial t} \ln f + f \frac{1}{f} \frac{\partial f}{\partial t} \right) d\vec{r} d\vec{p} \\ &= \int \frac{\partial f}{\partial t} \ln f d\vec{r} d\vec{p} + \frac{d}{dt} \int f d\vec{r} d\vec{p}, \end{aligned} \tag{3.6}$$

where the last term is 0 as it is a time derivative of a constant normalisation of f. For the first term, we can directly substitute in Eq. (3.2) after rearranging it for $\frac{\partial f}{\partial t}$. Ignoring the collision integral for now, we have the integrand simplify to

$$\begin{aligned} &\left(\frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} \cdot \frac{\partial f}{\partial \vec{r}} - \frac{\partial \mathcal{H}_{(1)}}{\partial \vec{r}} \cdot \frac{\partial f}{\partial \vec{p}} \right) \ln f \\ &= \left(\frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} \cdot \frac{\partial}{\partial \vec{r}} - \frac{\partial \mathcal{H}_{(1)}}{\partial \vec{r}} \cdot \frac{\partial}{\partial \vec{p}} \right) (f \ln f - f) \\ &= \frac{\partial}{\partial \vec{r}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} (f \ln f - f) \right) - \frac{\partial}{\partial \vec{p}} \cdot \left(\frac{\partial \mathcal{H}_{(1)}}{\partial \vec{r}} (f \ln f - f) \right) \end{aligned} \tag{3.7}$$

where we again used Theorem 1.3.1 and $\frac{\mathrm{d}}{\mathrm{d}f}(f\ln f - f) = \ln f$ (yes we are differentiating with respect to a function). These terms therefore integrate out to zero and we can move on.

Let us not forget our $C[f]$ still exists, and so the time derivative is

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \int \ln f(\vec{p_1}) \left[\int \left(f\left(\vec{p_1}'\right) f\left(\vec{p_2'}\right) - f(\vec{p_1}) f\left(\vec{p_2}\right) \right) W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) d\vec{p_2} d\vec{p_1}' d\vec{p_2'} \right] d\vec{r_1} d\vec{p_1} \tag{3.8}$$

Theorem 3.2.1.

$$\frac{\mathrm{d}H}{\mathrm{d}t} \leq 0 \iff \frac{\mathrm{d}S}{\mathrm{d}t} \geq 0 \tag{3.9}$$

Namely, entropy is always increasing or extremised. Equivalently, ***entropy can never decrease****.*

Proof. Time for a bit of algebra. The term $\ln f(\vec{p}) = \frac{1}{2}\ln(f(\vec{p_1})f(\vec{p_2}))$ remembering that we can exchange $\vec{p_1} \leftrightarrow \vec{p_2}$. Now remember by Theorem 3.1.1 we can exchange the primed and unprimed coordinates whenever we want. So therefore we can split the term again and $\frac{1}{2}\ln(f(\vec{p_1})f(\vec{p_2})) = \frac{1}{4}\left[\ln(f(\vec{p_1})f(\vec{p_2})) - \ln(f(\vec{p_1}')f(\vec{p_2}'))\right] = \frac{1}{4}\left[\ln(f(\vec{p_1})f(\vec{p_2})/f(\vec{p_1}')f(\vec{p_2}'))\right]$ by properties of logarithms.

We define

- $X = f(\vec{p_1})f(\vec{p_2})$
- $X' = f(\vec{p_1}')f(\vec{p_2}')$

We can substitute this for $\ln(f(\vec{p_1}))$ in Eq. (3.8) and obtain

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \int (X' - X)\ln\left(\frac{X}{X'}\right) W\left(\vec{p_1}', \vec{p_2}' \to \vec{p_1}, \vec{p_2}\right) d\vec{p_2} d\vec{p_1}' d\vec{p_2'} d\vec{r_1} d\vec{p_1} \leq 0 \tag{3.10}$$

This attains the first inequality. Simply substituting in the definition for H proves the $\iff$. Therefore entropy is always extremised or increasing - more specifically, **it can never decrease**. ☐

Corollary 3.2.1. *A system always tends towards a state of maximum entropy.*

Proof. We showed that entropy is always extremised. Moreover we showed it is never decreasing. Therefore if entropy achieves an extremum it must be a maximum. ☐

A system of particles interacting via some Hamiltonian obeys an N-particle Liouville equation (1.14). Its N-particle distribution function has zero total time derivative, and so entropy is conserved and it can never reach a state of maximum entropy.

Somewhere between introducing the N-particle Liouville theorem, and the Boltzmann equation which approximates the corresponding single particle equation in the BBGKY hierarchy, we've introduced the means for our system to increase its entropy. This happened when we made the assumption of molecular chaos (see Definition 3.3.1). Requiring that particle momenta decorrelate at some time after each collision has led to the situation that entropy can never decrease in our system, giving us a statistical description of particle motion consistent with the 2nd law of thermodynamics.

3.3 Equilibrium and the Maxwell Distribution

This section aims to prove one main theorem

Theorem 3.3.1. *The probability distribution of a system in thermal equilibrium is given by the Maxwell-Boltzmann distribution.*

$$f_{eq}(\vec{r},\vec{p}) = Ae^{-\mathcal{H}/T} = A\exp\left[\frac{-p^2/2m - \Phi_{ext} + \mu}{T}\right] \tag{3.11}$$

where T is temperature in units where Boltzmann's constant $k_B = 1$, μ is chemical potential, and A is some normalising constant such that f_{eq} integrates to N - see Eq. (1.10).

Before proving this theorem, we must define what equilibrium means.

Definition 3.3.1. An ensemble is in **equilibrium** if the following two conditions are satisfied:

- The distribution function $f(\vec{r},\vec{p},t)$ is stationary. We write $f_{\text{eq}}(\vec{r},\vec{p})$ with no explicit time dependence to be the stationary function.
- **Detailed balance** (*PX3A7 Statistical Physics*). This is the condition that the rate at which particles find their way from a point $(\vec{r},\vec{p})$ to some other point $(\vec{r'},\vec{p'})$ must be equal to the rate of the reverse process.

$$f_{\text{eq}}(\vec{p_1}')f_{eq}(\vec{p_2}') = f_{eq}(\vec{p_1})f_{eq}(\vec{p_2}) \tag{3.12}$$

Lemma 3.3.2. *The advection terms in Eq. (3.2) are zero for any function $f = g(\mathcal{H})$ that does not explicitly depend on $\vec{r}$ or $\vec{p}$.*

Proof. This is clear since the derivative terms $\dfrac{\partial f}{\partial \vec{r}}, \dfrac{\partial f}{\partial \vec{p}} = 0$, leaving only $\dfrac{\partial f}{\partial t}$ □

Corollary 3.3.1. *The functions satisfying Lemma 3.3.2 and Definition 3.3.1 (detailed balance) extremises entropy H.*

This is useful!

Proof. Substituting $f \to f_{\text{eq}}$ in Eq. (3.8) and using the detailed balance condition causes the integrand in Eq. (3.8) to be 0. By Corollary 3.2.1, systems obeying detailed balance are those which maximise entropy. □

We can express detailed balance as a conservation law by taking logarithms of both sides:

$$\ln f_{\text{eq}}(\vec{p_1}') + \ln f_{\text{eq}}(\vec{p_2}') = \ln f_{\text{eq}}(\vec{p_1}) + \ln f_{\text{eq}}(\vec{p_2}) \tag{3.13}$$

Therefore $\ln f_{\text{eq}}$ summed over the two particles involved, is conserved by elastic collisions. The only conserved quantities we actually have are energy and momentum, so $\ln f_{\text{eq}}$ must be a scalar additive affine function of those taking the form

$$\ln f_{\text{eq}}(\vec{p}) = \vec{\alpha}\cdot\vec{p} - \beta\mathcal{H} + \gamma \tag{3.14}$$

where each term is conserved by collisions when summed over the two colliding particles. Now we have two cases

Non uniform $\Phi_{\textbf{ext}}(\vec{r})$. Then we do not have conservation of momentum and must take $\alpha = 0$. This then gives the familiar **Boltzmann (thermal) distribution**

$$f_{\text{eq}} \propto \exp\left[\beta\mathcal{H}\right] \tag{3.15}$$

where the Maxwell-Boltzmann distribution Eq. (3.11) is a case.

Uniform $\Phi_{\textbf{ext}}(\vec{r})$. We do have conservation of momentum and hence non-zero α allowed, which turns out to correspond to thermal equilibrium in a moving frame with velocity $\vec{u}$.

Our single-particle Hamiltonian is

$$\mathcal{H} = \frac{p^2}{2m} + \Phi_{\text{ext}}$$

However the real kinetic energy is $\frac{1}{2m}\left|\vec{p} - m\vec{u}\right|^2 = \frac{p^2}{2m} - \vec{u}\cdot\vec{p} + \frac{m}{2}u^2$ so we identify $\vec{\alpha} = \vec{u}/k_B T$. Please remember so far we've been working with $k_B = 1$ but in SI units you need that constant!

Therefore the general equilibrium solution for this case is

$$\boxed{f_{\text{eq}} = f_0(n, T, \vec{u}) = n\left(2\pi m T\right)^{-3/2}\exp\left[-\frac{\left|\vec{p} - m\vec{u}\right|^2}{2mT}\right]} \tag{3.16}$$

where

- We substituted $\vec{\alpha}$ for its definition in terms of $\vec{u}$.
- $\beta = 1/T$, the **thermodynamic temperature**.
- γ and the uniform potential Φ_{ext} has been absorbed into a prefactor n
- $\left(2\pi m T\right)^{-3/2}$ is a normalising constant such that

  $$\int f_{\text{eq}} d\vec{p} = n(\vec{r})$$

  the real particle density as from Definition 1.5.1.

We can now check that this distribution is an equilibrium if Lemma 3.3.2 is satisfied.

Lemma 3.3.3. *f_0 defined above also satisfies Lemma 3.3.2, i.e. it gives zero in the advective terms of the Boltzmann equation*

Proof. Assuming $\vec{u}, T, n$ are independent of position, then yes as the divergence terms are again zero. ☐

So we have shown for reversible processes in equilibrium, you recover something which looks akin to a Boltzmann distribution and that the 2nd law of thermodynamics has been verified. Using this, we can consider what can happen as a system is only *locally* in equilibrium in the next chapter.

Chapter 4

Local Equilibrium and Transport

4.1 Local Equilibrium

We allow $\Phi_{\text{ext}}(\vec{r}), \vec{u}(\vec{r}), T(\vec{r}), n(\vec{r})$ to vary slowly with position (but not momentum).

Lemma 4.1.1. *With collisions viewed as point-like, the resulting local equilibrium*

$$f_0(n(\vec{r}), T(\vec{r}), \vec{u}(\vec{r}))$$

will obey detailed balance in the collision integral.

Proof. Indeed, whilst there is now explicit dependence on $\vec{r}$, the fact it is an equilibrium around some point $\vec{r_0}$ does mean that the advective terms in the Boltzmann equation are zero. The system *is effectively in* equilibrium in an open set $U = \{\vec{r} : |\vec{r} - \vec{r}_0| < |\delta\vec{r}|\}$. However outside of this range it can no longer be considered in equilibrium. □

Consider a ball kicked from the top of a hill into a valley on both sides with a trough at the bottom. The exact position of that trough is the site of total equilibrium, and perhaps some positions around the trough will be stable enough that the ball doesn't budge. In fact that trough is also a **stable equilibrium** since even a displacement will cause it to converge back to that point. However go far enough from the trough and the ball will no longer be in equilibrium.

Another local equilibrium is the perfect top of the hill, assuming there is a surface which it can stay still on. It is also locally at equilibrium here, however it is an **unstable equilibrium** as a small nudge of the ball is all it needs to start moving dramatically.

There are flaws with this analogy: a ball is a single particle, and gases aren't. The ball is non-interacting but particles in our system may interact with each other. Hopefully it just gets the point across about local equilibrium.

Anyways, we now want to see how the system behaves with this new equilibrium distribution. We want to handle the resulting new spatial gradients perturbatively, seeking a stationary solution of the form

$$f(\vec{r}, \vec{p}) = f_0(\vec{u}(\vec{r}), T(\vec{r}), n(\vec{r})) + \epsilon\delta f(\vec{r}, \vec{p})$$

with $\delta f \propto \vec{\nabla}\vec{u}, \vec{\nabla}T, \vec{\nabla}n, \vec{\nabla}\Phi_{ext}$, all assumed to be small gradients.

Remark. *Spatial gradients give current densities of momentum, heat and particles respectively.*

We now substitute this perturbation into the Boltzmann equation. First, we substitute this into $C[f]$, omitting the parameters for W since it is the same as before:

$$C[f] = \int W\left[f(\vec{p_1}'f(\vec{p_2}') - f(\vec{p_1})f(\vec{p_2})\right] \tag{4.1}$$

$$\begin{aligned} C[f+\delta f] = \int W\Big[&\left(f(\vec{p_1}') + \delta f(\vec{p_1}')\right)\left(f(\vec{p_2}') + \delta f(\vec{p_2}')\right) \\ &- \left(f(\vec{p_1} + \delta f(\vec{p_1})\right)\left(f(\vec{p_2}) + \delta f(\vec{p_2})\right)\Big] d\vec{p_2} d\vec{p_1}' d\vec{p'_2} \end{aligned} \tag{4.2}$$

Expanding all the brackets (it's a bit long but the algebra is simple just don't lose track of the indices), we can collect the like terms as so:

$$= \int W\Big[\left(f\left(\vec{p_1}'\right) f\left(\vec{p_2}'\right) - f\left(\vec{p_1}\right) f\left(\vec{p_2}\right)\right) \tag{4.3}$$

$$\begin{aligned} &+ \left(f\left(\vec{p_1}'\right) \delta f\left(\vec{p_2}'\right) - f\left(\vec{p_1}\right) \delta f\left(\vec{p_2}\right)\right) \\ &+ \left(f\left(\vec{p_2}'\right) \delta f(\vec{p_1}') - f\left(\vec{p_2}\right) \delta f\left(\vec{p_1}\right)\right) \end{aligned} \tag{4.4}$$

$$+ \left(\delta f\left(\vec{p_1}'\right) \delta\left(\vec{p_2}'\right) - \delta f\left(\vec{p_1}\right) \delta f\left(\vec{p_2}\right)\right)\Big] d\vec{p_2} d\vec{p_1}' d\vec{p'_2} \tag{4.5}$$

Notice in the square brackets we have terms of the form $ff - ff, f\delta f - f\delta f$ etc. for different permutations of $\vec{p_1}, \vec{p_2}$. We thus redefine $C[f]$ slightly (it is inherently a notation change but mathematically it must be clarified)

Definition 4.1.1. The Collision Integral is a bilinear functional $C : \mathbb{R}^2 \to \mathbb{R}$ of 2 functions $g, h : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ such that

$$C[g,h] = \int W[g(x')h(y') - g(x)h(y)]dydx'dy' \tag{4.6}$$

$$C[h,g] = \int W[h(x')g(y') - h(x)g(y)]dydx'dy' \tag{4.7}$$

We now must substitute our $\mathcal{H}$ into the advective terms of the Boltzmann equation. We know from Lemma 1.3.1 that any gradient terms of the Hamiltonian are zero, but due to the potential $\Phi_{\text{ext}}(\vec{r})$ having explicit dependence on $\vec{r}$, we keep $\vec{\nabla}\Phi_{\text{ext}}$ and we get

$$A[f] \coloneqq \frac{\vec{p}}{m} \cdot \frac{\partial f}{\partial \vec{r}} - \frac{\partial \Phi_{ext}}{\partial \vec{r}} \cdot \frac{\partial f}{\partial \vec{p}} \tag{4.8}$$

The Boltzmann equation is therefore

$$\frac{\partial f}{\partial t} + \epsilon A[f] = C[f,f] \tag{4.9}$$

Now substitute $f \to f_0 + \epsilon\delta f$ we get

$$\epsilon A[f_0] + O(\epsilon^2) = C[f_0, f_0] + \epsilon C[\delta f, f_0] + \epsilon C[f_0, \delta f] + O(\epsilon^2)$$

We know $C[f_0, f_0] = 0$ because it is a local-equilibrium distribution. Keeping only first-order terms in ϵ:

$$A[f_0] = C[\delta f, f_0] + C[f_0, \delta f] \coloneqq LC[\delta f] \tag{4.10}$$

This is an inhomogenous linear equation for δf, the **linearised collision integral** $LC[\delta f, f_0]$. Then we have $\delta f = LC^{-1}(A[f_0])$.

4.1.1 Advective term from f_0

Non-zero gradients of $n, T, u, \Phi_{ext} \implies A[f_0], \delta f \neq 0$. This leads to non-zero currents of particles, heat and momentum to drive the system to equilibrium. As from before,

$$LC[\delta f] = C[\delta f, f_0] + C[f_0, \delta f] = -\delta f/\tau. \tag{4.11}$$

i.e., scattering is *isotropic*.

We will define the **vector** $\vec{\nu} = \vec{\nabla}(\vec{p} - m\vec{u}(\vec{r})) = -m\vec{\nabla}\vec{u}$ because the spatial gradient is at constant $\vec{p}$. In the notes, he uses the vector $\vec{\pi}$. We avoid this convention. Then

$$f = f_0 + \delta f = f_0 \left(1 - \frac{\tau A[f_0]}{f_0}\right) \tag{4.12}$$

We now substitute Eq. (3.16) (general equilibrium solution) into Eq. (4.8) and evaluate the advective term of our new Boltzmann equation. We must differentiate carefully, and we make some assumptions as we derive. Note

$$\frac{A[f_0]}{f_0} = \frac{1}{f_0}\left[\frac{\vec{p}}{m} \cdot \vec{\nabla} f_0 - \vec{\nabla}\Phi_{\text{ext}} \cdot \vec{\nabla}_{\vec{p}} f_0\right] \tag{4.13}$$

where $\vec{\nabla} = \dfrac{\partial}{\partial \vec{r}}$ and $\vec{\nabla}_{\vec{p}} = \dfrac{\partial}{\partial \vec{p}}$

Lemma 4.1.2. *Prove*

$$\frac{\vec{\nabla} f_0}{f_0} = \frac{\vec{\nabla} n}{n} - \frac{3\vec{\nabla} T}{2T} - \vec{\nabla}\left(\frac{\nu^2}{2mT}\right) \tag{4.14}$$

Proof. This is a regular differentiation exercise. Remember that everything is a function of $\vec{r}$. We rewrite Eq. (3.16) as

$$\begin{aligned} f_0(n(\vec{r}), T(\vec{r}), \vec{u}(\vec{r})) &= n(\vec{r})\,(2\pi m T(\vec{r}))^{-3/2} \exp\left[-\frac{|\vec{\nu}|^2}{2mT(\vec{r})}\right] \\ &= n\Upsilon(T)\Theta(\vec{\nu}, T) \end{aligned} \tag{4.15}$$

Consider differentiation in one component $\dfrac{\partial}{\partial r_\mu}$. Therefore using the product rule twice

$$\frac{\partial f_0}{\partial r_\mu} = \frac{\partial n}{\partial r_\mu}\Upsilon\Theta + n\Theta\frac{\partial \Upsilon}{\partial r_\mu} + n\Upsilon\frac{\partial \Theta}{\partial r_\mu}$$

Note by the chain rule $\frac{\partial \Upsilon}{\partial r_\mu} = \frac{\partial \Upsilon}{\partial T}\frac{\partial T}{\partial r_\mu} = -\frac{3}{2}(2\pi m)^{-3/2}T^{-5/2}\frac{\partial T}{\partial r_\mu} = -\frac{3}{2}\frac{\Upsilon}{T}\frac{\partial T}{\partial r_\mu}$.

By the chain rule also, $\frac{\partial \Theta}{\partial r_\mu} = \Theta \frac{\partial}{\partial r_\mu}\left(\frac{\nu^2}{2mT}\right)$.

Now we can turn everything into vector form again by computing the derivative for each component. We are currently at

$$\vec{\nabla} f_0 = \Upsilon\Theta\vec{\nabla} n \times \frac{n}{n} - n\Theta\Upsilon\frac{3}{2}\frac{\vec{\nabla}T}{T} + n\Theta\Upsilon\vec{\nabla}\left(\frac{\nu^2}{2mT}\right) \tag{4.16}$$

Notice every term has a factor $f_0 = n\Upsilon\Theta$ so we factor these out and divide. This proves our lemma. □

We substitute this back into Eq. (4.13) and note that $\vec{\nabla}_{\vec{p}} f_0 = \frac{\vec{\nu}}{mT}$. Now substitute for every derivative term, allowing us to obtain

$$\frac{A[f_0]}{f_0} = \left(\frac{\vec{\nu}}{m} + \vec{u}\right)\cdot\left(\frac{\vec{\nabla}n}{n} + \left(\frac{\nu^2}{2mT} - \frac{3}{2}\right)\frac{\vec{\nabla}T}{T} + \frac{1}{T}(\vec{\nabla}\vec{u})\cdot\vec{\nu}\right) + \frac{\vec{\nu}}{mT}\cdot\vec{\nabla}\Phi_{\text{ext}}$$

- We ignore the extra $\vec{u}$ in the term $\left(\frac{\vec{\nu}}{m} + \vec{u}\right)$ whilst keeping its gradient $\vec{\nabla}\vec{u}$. This is because the direction $\vec{u}$ is parallel to the flow and changes in this direction are higher order contributions which we are neglecting.

Following this assumption through, then after some rearranging we obtain

$$\frac{A[f_0]}{f_0} = \frac{\vec{\nu}}{m}\cdot\left(\frac{\vec{\nabla}n}{n} + \left(\frac{\nu^2}{2mT} - \frac{3}{2}\right)\frac{\vec{\nabla}T}{T} + \frac{1}{T}\vec{\nabla}\Phi_{\text{ext}}\right) + \frac{1}{mT}\vec{\nu}\cdot(\vec{\nabla}\vec{u})\cdot\vec{\nu}. \tag{4.17}$$

4.1.2 Relaxation Time Approximation

Definition 4.1.2. The quantity τ is the **average time between scattering events**. It is the **relaxation time**.

This is the same τ met in Chapter 2.

Using the Mean Field, we can invoke our equations and note that $\tau \sim c/\lambda$ so

$$LC[\delta f] = -\frac{c\delta f}{\lambda} \tag{4.18}$$

This is the **Relaxation Time Approximation**.

Then the average with fixed temperature T gives $\bar{c}$ for **impurity scattering**. Additionally, for binary collisions, the above becomes

$$LC[\delta f] = -\frac{c_{\text{av}}\delta f}{\lambda} \tag{4.19}$$

where $c_{\text{av}} = \sqrt{2}\bar{c}$ is the **thermal average** (from the Maxwell distribution).

4.2 Impurity Scattering and Diffusion

The analysis in the previous section applies to microcanonical particles subject to impurity scattering. Consider the case where $\vec{u} = 0 \implies \vec{\nu} = \vec{p}, n = n(\vec{r})$ and $\vec{\nabla} T = 0$. Since $\vec{u}$ is zero everywhere, it follows $\vec{\nabla}\vec{u} = 0$

Then Eq. (4.13) reduces to

$$\frac{A[f_0]}{f_0} = \frac{\vec{p}}{m} \cdot \frac{\vec{\nabla} n}{n} + \frac{1}{T}\vec{\nabla}\Phi_{\text{ext}} \tag{4.20}$$

The number density still depends on position, so we have a *current* of particles. We can use Definition 1.5.2 to calculate the **response current density** at a point in space. We substitute $f \to f_0 + \delta f$ using Eq. (4.12) and the RTA. Then

$$\vec{j} = \int \frac{\vec{p}}{m}\frac{-\vec{p}}{m} \cdot \left(\frac{\vec{\nabla} n}{n} + \frac{1}{T}\vec{\nabla}\Phi_{\text{ext}}\right) \tau f_0 d\vec{p} \tag{4.21}$$

$$= \frac{-\tau}{m^2}\langle \vec{p}\vec{p}\rangle \cdot \left(\vec{\nabla} n + \frac{n}{T}\vec{\nabla}\Phi_{ext}\right) \tag{4.22}$$

where first term is expectation of outer product of momentum ($2m \times M_{ii}$, isotropic tensor). Notice we are integrating over $d\vec{p}$ so we get the spatial number density n from Definition 1.5.1 and so we can think of the product of momenta as an expectation value.

4.2.1 Second moments of the Maxwell Distribution

Definition 4.2.1. The **second-order moment tensor** $M_{\mu\nu}$ is defined as the product expectation value of some continuous (or discrete, but we state continous version here) random variable X with density $h(x)$:

$$M_{\mu\nu} = \int x_\mu x_\nu h(x) dx = \langle x_\mu x_\nu \rangle \tag{4.23}$$

Remark. *Consider the similarity between this and the stress tensor in Eq.* (1.5.3).

In our case, we have a distribution of momenta. We know that

$$f_0 \propto e^{-\frac{p^2}{2m}} = e^{-\frac{p_x^2}{2m}} e^{-\frac{p_y^2}{2m}} e^{-\frac{p_z^2}{2m}} \tag{4.24}$$

Then, applying standard integral results,

$$M_{ii} = \langle p_i^2 \rangle = \frac{\int p_i^2 e^{-p_i^2/2mT} dp_i}{\int e^{-p_i^2/2mT} dp_i} \implies q_i := \frac{p_i}{\sqrt{2mT}} \implies M_{ii} = mT$$

Lemma 4.2.1. *$M_{\mu\nu} = 0$ for $\mu \neq \nu$.*

Proof. This is clear by standard integral results. Using Definition 4.2.1 with $x_\mu x_\nu \to p_\mu p_\nu$ for $\mu \neq \nu$ and $f \to f_0$)as in Eq. (4.24)) we get that

$$\langle p_\mu p_\nu \rangle = \int p_\mu p_\nu \exp\left[\frac{p_\mu p_\mu}{2mT}\right] dp_\mu = 0$$

□

Physically, this means momenta in orthogonal directions are not correlated with each other. This makes sense because they are independent degrees of freedom.

Corollary 4.2.1.

$$M_{\mu\nu} = mT\delta_{\mu\nu} \equiv mT\mathbf{I}$$

where **I** *is the second-rank isotropic tensor - alternatively the Kronecker delta.*

Given this, we write $\vec{j}$ as

$$\vec{j} = -D\left(\vec{\nabla}n + \frac{n}{T}\vec{\nabla}\Phi_{ext}\right) \tag{4.25}$$

$$\implies D = \frac{\tau T}{m} = \frac{1}{3}\tau\overline{c^2} \tag{4.26}$$

4.2.2 Einstein Relation

Now the external potential induces a force $\vec{F} = -\vec{\nabla}\Phi_{\text{ext}}$. This causes an acceleration, and therefore a velocity response

$$\vec{v} = \frac{\vec{F}}{\alpha}$$

where α is the **drag coefficient**.

The current density $\vec{j} = n\vec{v} = -\frac{n}{\alpha}\vec{\nabla}\Phi_{\text{ext}}$. Compare the coefficient to above, we have

$$-\frac{Dn}{T}\vec{\nabla}\Phi_{\text{ext}} = -\frac{n}{\alpha}\vec{\nabla}\Phi_{\text{ext}} \implies D = \frac{k_B T}{\alpha}$$

where $k_B = 1$ in our units but we rewrite it for SI.

4.2.3 Quasi-Static Evolution

In the limit where spatial gradients are small and evolution is slow (i.e. requires large timesteps) we can use a **quasi-static evolution** where we set $\vec{j} = \vec{j}_{\text{static}}$. By Eq. 1 we have

$$\frac{\partial n}{\partial t} = -\vec{\nabla}\cdot\vec{j} = \vec{\nabla}\cdot -D\left(\vec{\nabla}n + \frac{n}{T}\vec{\nabla}\Phi_{ext}\right) \tag{4.27}$$

This has mathematical and physical consequences. Ignoring the potential term, we have the PDE

$$\frac{\partial n}{\partial t} = -\vec{\nabla}\cdot\vec{j} = -D\left(\vec{\nabla}^2 n\right) \tag{4.28}$$

This is the **Diffusion Equation**. Introducing the potential term as before is an example of a **Fokker-Planck equation**.

Remark. *Fokker-Planck equations are great. You can use them to study diffusion with external or random forces, such as in Brownian motion.*

4.3 Stress and Viscosity

On some surface $d\vec{S}$, we have a force $d\vec{F}$ applied. Then $d\vec{F} = \Sigma d\vec{S}$. Alternatively, $dF_\alpha = \Sigma_{\alpha\beta} dS_\beta$. But since we apply the force, this is a rate of *injection of momentum*. We sub in $\vec{p} = m\vec{u} + \vec{\nu}$, with $\vec{\nu}$ from the comoving pressure.

Definition 4.3.1. The **pressure tensor** $\Sigma_{\mu\nu}$ is defined as

$$-\Sigma = \int \frac{\vec{p}\vec{p}}{m} f d\vec{p} \tag{4.29}$$

The second moment of f (full distribution of all particles) is the first interesting one: it tells us (minus) the stress tensor as the flux density of momentum. We do minus since we want the posiitve direction to be *injecting* momentum to the system.

$$-\Sigma = \int \frac{\vec{p}\vec{p}}{m} f d\vec{p} = \rho\vec{u}\vec{u} + \int (\vec{u}\vec{\nu} + \vec{\nu}\vec{u}) f d\vec{p} + \int \frac{\vec{\nu}\vec{\nu}}{m} f d\vec{\nu}$$

where $\rho = mn$ is the **mass density** and $\rho\vec{u}\vec{u}$ is the **Reynolds stress**.

Lemma 4.3.1.

$$\int (\vec{u}\vec{\nu} + \vec{\nu}\vec{u}) f d\vec{p} = -\tau\vec{u} \cdot \vec{\nabla}(\rho\vec{u}\vec{u})$$

This term is the gradient correction to the Reynolds stress.

Proof. Non-examinable. □

Considering we still have uniform T, Φ_{ext}, n then substituting $f = f_0 + \delta f$ gives the last term as

Lemma 4.3.2. *Prove that the last term is a momentum average ($\vec{\nu}$) over f_0. Namely,*

$$\int \frac{\vec{\nu}\vec{\nu}}{m} f_0 \left(1 - \tau A\left[f_0\right]/f_0\right) d\vec{\nu} = n \left\langle \frac{\vec{\nu}\vec{\nu}}{m} \left(1 - \frac{\tau}{mT} \vec{\nu} \cdot (\vec{\nabla}\vec{u}) \cdot \vec{\nu}\right) \right\rangle_0$$

where $\langle h(\nu) \rangle_0 = \int h(\nu) f_0 d\vec{\nu}$, the momentum expectation value.

Proof. So it is easier to see, we start from the LHS. Note, from Eq. (4.15) that $f_0 \propto n(\vec{r}) \exp\left[-\nu^2/2mT\right]$. We will substitute in Eq. (4.17) with the uniform gradients and get

$$\int \frac{\vec{\nu}\vec{\nu}}{m} f_0 \left(1 - \frac{\tau A\left[f_0\right]}{f_0}\right) d\vec{\nu} = \int f_0 \underbrace{\frac{\vec{\nu}\vec{\nu}}{m} \left(1 - \frac{\tau}{mT} \vec{\nu} \cdot (\vec{\nabla}\vec{u}) \cdot \vec{\nu}\right)}_{h(\nu)} d\vec{\nu} = n \int f_0 h(\vec{\nu}) d\vec{\nu}$$

We have suggestively written this as an expectation value and factored out n from f_0 so there is again an abuse of notation. Therefore writing the last step we see

$$n \int f_0 h(\vec{\nu}) d\vec{\nu} = n \left\langle \frac{\vec{\nu}\vec{\nu}}{m} \left(1 - \frac{\tau}{mT} \vec{\nu} \cdot (\vec{\nabla}\vec{u}) \cdot \vec{\nu}\right) \right\rangle_0 \tag{4.30}$$

□

with $\vec{\nabla}\vec{u}$ the **shear stress tensor**. In particular, distributing the expectation value to products we see that this equals to

$$\frac{n}{m}\langle\vec{\nu}\vec{\nu}\rangle_0 - \frac{n\tau}{m^2T}\langle\vec{\nu}\vec{\nu}\vec{\nu}\cdot(\vec{\nabla}\vec{u})\cdot\vec{\nu}\rangle_0 \tag{4.31}$$

Lemma 4.3.3. *Prove*

$$\langle\nu_\alpha\nu_\beta\nu_\gamma\nu_\delta\rangle = (mT)^2\left(\delta_{\alpha\beta}\delta_{\gamma\delta}+\delta_{\alpha\gamma}\delta_{\beta\delta}+\delta_{\alpha\delta}\delta_{\gamma\beta}\right) \tag{4.32}$$

Proof. Non-examinable, however we will need this. For a general proof before taking expectation values, see Hodge, 1961. The general idea is that the general 4th rank isotropic tensor can be decomposed as a sum of tensor products of 2nd rank isotropic tensors - i.e. Kronecker deltas. The general expression is

$$\nu_\alpha\nu_\beta\nu_\gamma\nu_\delta = \left(\mathcal{A}\delta_{\alpha\beta}\delta_{\gamma\delta}+\mathcal{B}\delta_{\alpha\gamma}\delta_{\beta\delta}+\mathcal{C}\delta_{\alpha\delta}\delta_{\gamma\beta}\right) \tag{4.33}$$

with $\mathcal{A},\mathcal{B},\mathcal{C}$ constants specific to the problem.

By Corollary (4.2.1), we see that taking expectation values means $\mathcal{A},\mathcal{B},\mathcal{C}=(mT)^2$. This concludes the proof. □

Now we return back to Eq. (4.3.1) and substitute in for all the values we have so far. Then

Pressure Tensor

Theorem 4.3.4.

$$\Sigma = -\rho\vec{u}\vec{u} - nT\mathbf{I} + nT\tau\left(\vec{\nabla}\vec{u}+(\vec{\nabla}\vec{u})^T+\mathbf{I}\vec{\nabla}\cdot\vec{u}\right) \tag{4.34}$$

where we identify

- $$\Sigma_{\text{Reynolds}} = \vec{\nabla}\vec{u}+(\vec{\nabla}\vec{u})^T$$
  as the ***viscous stress*** *or Reynolds stress tensor*
- $$\Sigma_{\text{bulk}} = \mathbf{I}\vec{\nabla}\cdot\vec{u}$$
  as part of the ***bulk viscosity***

Proof. NON-EXAMINABLE (probably, I mean this stuff is long...). There is a partial proof actually in the LaTeXcode for this but it's incomplete and about an entire page... TL;DR is to work through the outer products and indices and use the fact there is no convective acceleration to remove the term in $\vec{\nabla}(\rho\vec{u}\vec{u})$ term.

We obtain the following by dimensional homogeneity (SI units):

- The **ideal gas pressure** $P = nk_BT$
- The **viscosity** $\eta = nk_BT\tau$

□

4.3.1 Quasi-Static Evolution

Want to use continuity of momentum:

$$\frac{\partial}{\partial t}(nm\vec{u}) + \vec{\nabla}\cdot(-\Sigma) = 0$$

Typically we rewrite it with $\rho = mn$ and substituting in for Σ we get

$$\underbrace{\frac{\partial}{\partial t}(\rho\vec{u}) - \vec{\nabla}\cdot\Sigma_{\text{Reynolds}}}_{\rho\frac{D}{Dt}\vec{u}} = \vec{\nabla}\cdot\Sigma_{\text{viscous}} - \vec{\nabla}P$$

This is the **Navier-Stokes equation**.

4.4 Thermal Conductivity

We want to use Definition 1.5.4 to get the heat current density so

$$\vec{J} = \int \frac{\nu^2}{2m}\frac{\vec{p}}{m} f \, d\vec{p}$$

Consider a stationary fluid, so $\vec{\nu} = \vec{p}$. We need to track $\vec{\nabla}T, \vec{\nabla}n$ whilst we choose $\vec{\nabla}\Phi_{ext} = 0 \implies \vec{\nabla}P = 0$. From the ideal gas law, with $k_B = 1$, we have

$$\frac{\vec{\nabla}T}{T} = -\frac{\vec{\nabla}n}{n}$$

$$\implies \vec{J} = \int \frac{1}{2m^2} p^2 \vec{p} f_0 \left(1 - \tau A[f_0]/f_0\right) d\vec{\nu} + \mathcal{O}(\vec{u})$$

$$= -\frac{n\tau}{2m^2}\left\langle \nu^2 \vec{\nu} A\left[f_0\right]/f_0\right\rangle_0 \tag{4.35}$$

$$= -\frac{n\tau}{2m^2}\left(\left\langle \nu^2\vec{\nu}\frac{\vec{\nu}}{m}\right\rangle_0 \cdot \left(\frac{\vec{\nabla}n}{n} - \frac{3}{2}\frac{\vec{\nabla}T}{T}\right) + \left\langle \nu^2\vec{\nu}\frac{\vec{\nu}}{m}\nu^2\right\rangle_0 \frac{1}{2mT}\frac{\vec{\nabla}T}{T}\right) \tag{4.36}$$

$$= -\frac{n\tau T^2}{2m}\left(5\left(\frac{\vec{\nabla}n}{n}\right) + 10\frac{\vec{\nabla}T}{T}\right) \tag{4.37}$$

from substituting Eq. (4.13) into the heat current density. To get to the last line, we used extensions of previous lemmas to show

$$\langle \pi^2 \pi_\alpha \pi_\beta \rangle_0 = 5(mT)^2 \delta_{\alpha\beta} \tag{4.38}$$

$$\langle \pi^4 \pi_\alpha \pi_\beta \rangle_0 = 35(mT)^3 \delta_{\alpha\beta} \tag{4.39}$$

$$\vec{J} = -\frac{\tau T}{2m}\left(5\left(T\vec{\nabla}n\right) + 10n\vec{\nabla}T\right) = -\frac{5\tau T}{2m}\left(n\vec{\nabla}T\right) \tag{4.40}$$

We have derived the **thermal conductivity law**

$$\vec{J} = -K\vec{\nabla}T \tag{4.41}$$

Definition 4.4.1. The **thermal conductivity** is

$$K = -\frac{5\tau T n}{2m} \tag{4.42}$$

Definition 4.4.2. The **isobaric heat capacity** for an ideal monatomic gas is

$$C_p = \frac{5}{2}k_B \tag{4.43}$$

Remark. *Observe* $\tau T n = \eta$*, the viscosity so* $K = \frac{C_p \eta}{m}$

Definition 4.4.3. The **Prandtl number** defined as

$$\frac{\eta/m}{K/C_p} = 1 \tag{4.44}$$

We observe in experiments and simulations that for simple molecular gases such as air, it is about 0.7. This demonstrates that we are very much in an approximation regime.

4.5 Shortcomings

To improve our analysis, we can relax some assumptions made.

- Do not assume that scattering is isotropic as in Eq. (4.11).
- Velocity dependence of the collision rate τ, such as in Eq. (2.11) and Definition 4.1.2.
- No best choice of moments to take - do we always carry averages forward, or is there a point where we can stop using averages etc. Simplifying using averages (moments) introduce assumptions about the system behaving fairly uniformly.

Chapter 5

Diffusion

5.1 Quantum Fermi Gas

We go for a semiclassical treatment.

Definition 5.1.1. The **Pauli Exclusion Principle** (PEP), is that two particles cannot occupy the same single particle quantum state (or level).

This is coupled to how we count single particle states in free space.

Lemma 5.1.1. *For each of spin up* $(+)$ *and spin down* $(-)$ *the count of single particle states in phase space is given by*

$$dQ_\pm = \frac{V}{h^3} d\vec{p}$$

Proof. Deriving this should be familiar from PX285 Statistical Mechanics or PX262 Quantum Mechanics. Consider a cube with side lengths L of fermions in 3D. The volume $V \sim L^3$ for a small enough box. The Time-Independent 3D Schrödinger equation reads

$$E\vec{\psi} = -\frac{\hbar^2}{2m}\vec{\nabla}^2\vec{\psi} + V(\vec{r})\vec{\psi} \tag{5.1}$$

For a 3D box, we obtain the solution

$$\vec{\psi} \propto \begin{bmatrix} \sin(k_x x) \\ \sin(k_y y) \\ \sin(k_z z) \end{bmatrix}$$

where $k_\mu = \frac{2\pi}{L} n_\mu$ with n_μ the number density in direction μ.

Now $dQ = dn_x dn_y dn_z \implies dk = \frac{8\pi^3}{L^3} dQ$. Then as $\vec{p} = \hbar\vec{k} \implies dp = \hbar dk \implies dp = \frac{8\pi^3\hbar}{V} dQ$. Remember that $\hbar = \frac{h}{2\pi}$. Substituting this in proves the lemma. □

Corollary 5.1.1. *The* ***density of single-particle levels in phase space*** *is*

$$q_\pm = \frac{dQ_\pm}{V dp} = \frac{1}{h^3} \tag{5.2}$$

Remark. *In the notes, he uses $d\vec{p}$ but to keep the mathematics consistent, it's really a* ***phase space volume*** *d^3p we are talking about.*

Additionally $q_\pm = 1$ in natural units where $h = 1$.

The PEP forces states below the Fermi energy E_F to fill. In particular, in an isotropic system, the Fermi surface is a sphere[1] with radius equal to $p_F = \sqrt{p_\mu p_\mu}$ and $E_F = \frac{p_F^2}{2m} \gg k_B T$.

Definition 5.1.2. The **filling factor** is defined as

$$\frac{\text{phase space density of particles}}{\text{phase space density of states}} \tag{5.3}$$

Then the total number of particles $N = Q_+ + Q_- = 2\frac{V\frac{4\pi}{3}p_F^3}{h^3}$. Then

$$2mE_F = \left(\frac{3}{8\pi}nh^3\right)^{2/3} \tag{5.4}$$

5.1.1 Scattering and Equilibrium

Recall that f is a probability distribution for indistinguishable particles - fermions as we are currently considering - to take a momentum $\vec{p}$. The effect of the PEP is to reduce any scattering rate by a factor $(1-f') \coloneqq 1 - f(\vec{p_1}')$ for a scatter which takes a fermion from momentum $\vec{p_1} \to \vec{p_1}'$ which is the probability that the state scattered into is **unfilled**.

Since we have fermions, we have to consider different cases:

Reversible Impurity Scattering Unchanged net collision rates. The scattering (gain and loss functions) for a fermion $\vec{p_1} \to \vec{p_1}'$ is slightly different to Eqs. (2.7), namely the integrands are

$$W(\vec{p_1} \to \vec{p_2})f(\vec{p_1})(1 - f(\vec{p_1}')) \tag{5.5}$$

However, considering $S_{\text{in}} - S_{\text{out}}$ gives the integrands as

$$W(\ldots)(f(\vec{p_1})(1 - f(\vec{p_1}')) - f(\vec{p_1}')(1 - f(\vec{p_1}))) = W(\ldots)(f' - f)$$

Binary collisions The modification to probabilities do not all cancel. We must consider the probabilities of each fermion initially with momenta $\vec{p_1}, \vec{p_2}$. The scattering (gain and loss functions) for these fermions are again similar to those in the integrands of Eqs. (2.7):

$$\begin{aligned} &W\left(\vec{p_1}\vec{p_2} \to \vec{p_1}', \vec{p_2}'\right) \\ &\qquad \times \Big[f(\vec{p_1})(1 - f(\vec{p_1}'))f(\vec{p_2})(1 - f(\vec{p_2}')) \\ &\qquad - f(\vec{p_1}')(1 - f(\vec{p_1}))f(\vec{p_2}')(1 - f(\vec{p_2}))\Big] \end{aligned} \tag{5.6}$$

---

[1] Specifically, we have 3 momentum directions, so the Fermi surface is a **2-sphere** *embedded* in $\mathbb{R}^3$.

Our detailed balance in Definition 3.3.1 must now satisfy

$$f_{\text{eq}}(\vec{p_1})(1-f_{\text{eq}}(\vec{p_1}'))f_{\text{eq}}(\vec{p_2})(1-f_{\text{eq}}(\vec{p_2}'))=f_{\text{eq}}(\vec{p_1}')(1-f_{\text{eq}}(\vec{p_1}))f_{\text{eq}}(\vec{p_2}')(1-f_{\text{eq}}(\vec{p_2}))$$

This equation can be rearranged into the form

$$\frac{f_1'}{(1-f_1')}\frac{f_2'}{(1-f_2')}=\frac{f_1}{(1-f_1)}\frac{f_2}{(1-f_2)}$$

with the primed and unprimed coordinates on separate sides. Again, we rewrite this as a conservation law by taking logarithms of both sides

$$\log\left(\frac{f_{\text{eq}}\ (\vec{p_1}')}{1-f_{\text{eq}}\ (\vec{p_1}')}\right)+\log\left(\frac{f_{\text{eq}}\ (\vec{p_2}')}{1-f_{\text{eq}}\ (\vec{p_2}')}\right)=\log\left(\frac{f_{\text{eq}}(\vec{p_1})}{1-f_{\text{eq}}(\vec{p_1})}\right)+\log\left(\frac{f_{\text{eq}}\ (\vec{p_2})}{1-f_{\text{eq}}\ (\vec{p_2})}\right) \tag{5.7}$$

Following the same idea by relating each log to some function of energy as in Eq. (3.14) we have

$$\frac{f_{\text{eq}}\ (\vec{p})}{1-f_{\text{eq}}\ (\vec{p})}\propto e^{-\beta(E-\mu)-\vec{\alpha}\cdot\vec{p}} \tag{5.8}$$

As in Section 3.3, we make no assumptions about the momentum as the momentum term can be chosen arbitrarily by changing frame of reference, so $\vec{\alpha}=0$ and we recover:

Fermi-Dirac Distribution

Theorem 5.1.2. *The Fermi-Dirac Distribution is the equilibrium distribution for an ensemble of fermions.*

$$f_{eq}=f_0(E)=\frac{1}{e^{\beta(E-\mu)}+1} \tag{5.9}$$

where

- *μ is the* ***chemical potential*** *of the system which at $T=0$ is the Fermi energy.*
- *$\beta=1/k_BT$ is the thermodynamic temperature yet again.*

This depends on $\vec{r}$ and $\vec{p}$ only through $\mathcal{H}$ (whose value is E), the advective terms of the Boltzmann equation also give zero with this distribution so we do have a full equilibrium:

$$f(\vec{r},\vec{p})=\frac{1}{e^{\beta(\mathcal{H}_{(1)}(\vec{r},\vec{p})-\mu(\vec{r}))}+1}$$

Remark.

- *μ vs E_F: As $T\to 0(\beta\to\infty),\mu=E_F$*
- *For $T>0$, μ is well-defined but E_F is not. The Fermi surface is fuzzy on scale of k_BT, i.e. a region around the Fermi momentum/energy where the occupation is not strictly 1 or 0.*

Theorem 5.1.3. *Using the same analysis, show that for bosons, we get the* ***Bose-Einstein Distribution****.*

$$f_{eq}=f_0(E)=\frac{1}{e^{\beta(E-\mu)}-1} \tag{5.10}$$

Proof. It is almost identical proof, however bosons do not obey PEP, so we can have multiple bosons in the same state, so the probability of entering a state is $1+f$. □

Remark. *Because these distributions depends on $\vec{r}$ and $\vec{p}$ only through $\mathcal{H}$ (whose value is E), the advective terms of the Boltzmann equation also give zero with this distribution so we do have a full equilibrium.*

5.1.2 Diffusion, Conductance and Einstein's Law

As for classical, binary scattering is crucial in determining the equilibrium energy distribution of the particles, but it makes no direct contribution to attenuating the particle current (density) because the sum of particle momenta and hence total current is conserved in binary collisions.

For elastic impurity collisions, energy is conserved at the single particle level and cannot adjust its distribution, but it certainly does attenuate the particle current.

We consider an ensemble obeying a FD distribution in Eq. (5.9). We suppose $\mu = \mu(\vec{r})$ a property of the system and an external potential $\Phi_{\text{ext}}(\vec{r})$ and their gradients, which we will say are non-zero. We seek to find the corresponding steady perturbation to the particle distribution and its current. Our perturbation is therefore

$$f = f_0(\mu(\vec{r})) + \delta f \tag{5.11}$$

Substitute this into the Boltzmann equation in Eq. (2.9) and we get

$$\frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} \cdot \vec{\nabla}\mu \frac{\partial f_0}{\partial \mu} - \vec{\nabla}\Phi_{\text{ext}} \cdot \frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} \frac{\partial f_0}{\partial E} = -\frac{\delta f}{\tau} \tag{5.12}$$

where we have introduced again the RTA as in Definition 4.1.2 for $C[f]$. We are only taking f_0 to first order $f_0 \approx \partial_\alpha f_0 \vec{\nabla}\mu$ since f_0 varies slowly and δf is a small correction and varying slowly.

Lemma 5.1.4. *Prove that the change in equilibrium with the chemical potential is the negative of the change in equilibrium with energy*

$$\frac{\partial f_0}{\partial \mu} = -\frac{\partial f_0}{\partial E} \tag{5.13}$$

Proof. We consider the change of variables $\mu(\vec{r}) \to E(\vec{r}) - \mu(\vec{r})$ from the FD distribution. Then $\dfrac{\partial E}{\partial \mu} = -1$. Substituting this using chain rule proves this. □

Then the Boltzmann equation simplifies to

$$\delta f = \tau \left(\vec{\nabla}\mu(\vec{r}) + \vec{\nabla}\Phi_{\text{ext}}(\vec{r}) \right) \cdot \frac{\partial \mathcal{H}_{(1)}}{\partial \vec{p}} \frac{\partial f_0}{\partial E} \tag{5.14}$$

Physical explanation: setting up a gradient of chemical potential and applying an external force field **have the same effect**. δf is concentrated near the Fermi surface because of the factor $\frac{\partial f_0}{\partial E} = -f_0(E)(1 - f_0(E))/T$ which becomes exponentially small for $|E - E_F|$ large compared to $k_B T$. In the quantum case all the scattering action and effect is concentrated around the Fermi surface.

We can interpret the result for δf in Eq. (5.14) as a Taylor expansion of the equilibrium distribution f_0 rigidly displaced in momentum.

Using the chain rule, we say $f_0(\mathcal{H}_{(1)})$ such that the product of derivatives simplifies to $\partial_{\vec{p}} f_0$, then leads to:

$$f(\vec{p}) = f_0(\mathcal{H}_{(1)}(\vec{p})) + \delta f = f_0(\vec{p} - \Delta\vec{p} + \mathcal{O}(\nabla^2))$$

where the last term contains all higher-order gradients which we will neglect and we define

Definition 5.1.3. The **momentum displacement** is the momentum acquired in a gain τ:

$$\Delta\vec{p} = -\tau(\vec{\nabla}\mu + \vec{\nabla}\Phi_{\text{ext}}) \tag{5.15}$$

Ignoring higher-order terms, we see the FD distribution is now centred on the momentum displacement. This is equivalent to a force acting on all the states on the Fermi surface for a time τ, which shifts the surface (*PX385 Condensed Matter Physics*).

Evaluating the current by considering $\Delta\vec{p}/m$ (5.14):

$$\vec{j} = n\Delta\vec{p}/m = -M(\vec{\nabla}\mu + \vec{\nabla}\Phi_{\text{ext}})$$

Definition 5.1.4. The quantity M is called the **mobility** defined as

$$M = \frac{n\tau}{m} \tag{5.16}$$

This is the ratio of the particle's velocity to the force $-\vec{\nabla}\Phi_{\text{ext}}$ applied.

We see there are 2 current-contributing terms: $-M\vec{\nabla}\mu$ and $-M\vec{\nabla}\Phi_{\text{ext}}$. In particular,

$$j = -M\vec{\nabla}\mu = -M\frac{\partial\mu}{\partial n}\vec{\nabla}n = -D\vec{\nabla}n \tag{5.17}$$

$$\tag{5.18}$$

We arrive at

Generral Einstein Relation

$$D = M\frac{\partial\mu}{\partial n} \tag{5.19}$$

This is the general form of the **Einstein relation**. In the classical limit, $M = nD/T$. The most relevant external driving force is electric field $\vec{E}$ on particles of charge $q : \vec{\nabla}\Phi_{\text{ext}} = q\vec{E}$ and:

$$\vec{j}_e = \sigma\vec{E} = (-M\vec{\nabla}\Phi_{\text{ext}})\vec{E} = \frac{nq^2\tau}{m}\vec{E}$$

Lemma 5.1.5.

$$\sigma = \frac{nq^2\tau}{m}$$

*This is **Drude's result** (PX385 Condensed Matter Physics and this Cambridge page for more information).*

Proof. This proof is non-examinable. □

In the diffusion equation, $0 = \vec{\nabla} \cdot \vec{j} + \partial_t n$. We can put expression for $j = -D\vec{\nabla}n$ inside:

$$\frac{\partial n}{\partial t} = D\vec{\nabla}^2 n \tag{5.20}$$

This is the **Diffusion equation**.

5.2 Weak Scattering, Momentum diffusion and Fokker-Planck Equation

5.2.1 Review of the Diffusion Equation

Definition 5.2.1. Fick's Law

$$\vec{j} = -D\vec{\nabla}n$$

Recall for classical impurity scattering, Eqs (4.25) gives

$$D = \frac{1}{3}\tau\overline{c^2} = \frac{1}{3}\lambda\bar{c} \tag{5.21}$$

and additionally, Eq. (5.19) for quantum scattering. Whatever one we use we get the diffusion equation Eq. (5.20).

Now, imagine particles on a line either in real space or momentum space. These particles jump/move a distance a at a rate $1/\tau$ but choose a direction randomly (step distribution). Define $n(x,t)$ be number between $U_{x,a} = [x - a/2, x + a/2]$ with properties directly above about rate and direction. Then

Lemma 5.2.1.

$$n(x, t + \delta t) = n(x,t) - n(x,t)\frac{\delta t}{\tau} + \frac{n(x+a,t)}{2}\frac{\delta t}{\tau} + \frac{n(x-a,t)}{2}\frac{\delta t}{\tau} \tag{5.22}$$

Proof. In a time δt, any 3 things can happen. The first term $n(x,t)$ is simply just the initial number of particles at time t. Since it is a count, it is nonnegative, $n(x,t) \geq 0$.

1. The term $n(x,t)\frac{\delta t}{\tau}$ is the number of particles as a fraction of $n(x,t)$ which choose to leave the interval $U_{x,a}$. This decreases the number of particles, hence the $-$ sign in front.

2. The term $\frac{n(x+a,t)}{2}\frac{\delta t}{\tau}$ describes the number of particles which choose to jump into $U_{x,a}$ from the right, i.e. interval $U_{x+a,a}$, as a fraction of the number of particles on the RHS which is $n(x+a,t)$. This increases the number of particles in $U_{x,a}$ so it is a positive quantity.

3. Similarly, the term $\frac{n(x-a,t)}{2}\frac{\delta t}{\tau}$ describes the number of particles which choose to jump into $U_{x,a}$ from the left, i.e. interval $U_{x-a,a}$, as a fraction of the number of particles on the LHS which is $n(x-a,t)$. This again would increase the number of particles in $U_{x,a}$

□

We perform a Taylor expansion and stop at second order assuming $n(x,t)$ varies slowly and is smooth over the interval $(x-a, x+a)$:

$$n(x \pm a, t) \approx n(x,t) \pm a\frac{\partial n}{\partial x} + \frac{a^2}{2}\frac{\partial^2 n}{\partial x\, \partial x} \tag{5.23}$$

Using the definition of a partial derivative, taking limits to get the time derivative in Eq. (5.22) and the spatial Taylor expansion Eq. (5.23) we see

$$\frac{\partial n}{\partial t} = \lim_{\delta t \to 0} = \frac{n(x,t+\delta t) - n(x,t)}{\delta t} = -\frac{a^2}{2\tau}\frac{\partial^2 n}{\partial x^2} \tag{5.24}$$

Remark.
- *Tells us*

$$D = \frac{a^2}{2\tau}$$

, with a the mean free path and τ is time between scattering events.

- *In 3D,*

$$D = a^2/6\tau$$

- *Diffusion equation models heat, diffusion (duh), random walks, Brownian motion, chemical reactions, financial systems and many more physical processes.*

- *The PDE is 2nd-order, homogeneous and linear.*

Suppose we add an **advective term** in Eq. (5.23) of the form

$$v\frac{\partial n}{\partial x}$$

where v is a speed. This acts to move the particles towards the right (or left if $v < 0$). We will show the following theorem

Theorem 5.2.2. *The addition of any advective terms does not affect the underlying physics of the particles, i.e. the particles still diffuse the same way.*

Mathematically, this is saying a linear transformation of variables of the advective diffusion equation will recover the regular diffusion equation in terms of the transformed variables.

Proof. Mathematical remark: we ignore the boundary conditions - these depend on the physical problem at hand and will be in terms of x, t. During the transformation, these boundary conditions *must also be transformed* into the new variables to be consistent. It is important that the transformation chosen preserves all the properties of the system.

We introduce a speed v to all the particles. Assuming $v \ll c$ we can perform a Galiliean transformation

$$t' = t, x' = x - vt$$

Using the multivariable chain rule, the derivatives become

$$\begin{aligned} \partial_x = \partial_x x' \partial_{x'} + \partial_x t' \partial_{t'} = \partial_{x'} \\ \partial_t = \partial_t x' \partial_{x'} + \partial_t t' \partial_{t'} = \partial_{x'} = -v\partial_{x'} + \partial_{t'} \\ \implies -v\partial_{x'} n + \partial_{t'} n + v\partial_{x'} n - D\partial_{xx} n = 0 \\ \implies D\frac{\partial^2 n}{\partial x'^2} = \frac{\partial n}{\partial t'} \end{aligned} \tag{5.25}$$

i.e. the advection terms vanish and we recover the original diffusion equation in (x', t') □

Intuitively, this makes sense. If we give every particle the same uniform boost in speed, all we need to do is change the reference frame to a frame which moves with speed v in the same direction of motion, so therefore the particles are once again diffusing as if there was no advection.

It is now time to move on to **solving this PDE**, you may have seen this in *Mathematics for Physicists* or *MA250 Introduction to PDEs.*

Theorem 5.2.3. *The* ***fundamental solution or Green function to the diffusion equation*** *Eq. (5.20).*

$$n(x,t) \sim \frac{1}{\sqrt{2\pi D(t-t_0)}} e^{-x^2/4D(t-t_0)} \tag{5.26}$$

for diffusion that starts at $t = t_0$, and particles which don't have a velocity $\vec{p}$.

Proof. We take a **spatial Fourier transform**

$$\tilde{n}(k,t) = \mathfrak{F}[n(x,t)] = \frac{1}{\sqrt{2\pi}} \int_\Gamma n(x,t) e^{ikx} dx \tag{5.27}$$

and apply this to the diffusion equation. Note that $\mathfrak{F}[\partial_t n] = \partial_t \tilde{n}$ since we assume n is smooth. Also, $\mathfrak{F}[\partial_{xx} n] = -k^2 \tilde{n}$. The diffusion equation now reads

$$\frac{\partial \tilde{n}}{\partial t} = -D\tilde{n} \tag{5.28}$$

We can directly solve this, namely

$$\tilde{n}(k,t) \propto e^{-Dk^2 t} \tag{5.29}$$

Now take the inverse spatial Fourier transform to get back to x:

$$\begin{aligned} n(k,t) = \mathfrak{F}^{-1}[\tilde{n}(k,t)] &= \frac{1}{\sqrt{2\pi}} \int_{\Gamma_k} \tilde{n}(k,t) e^{-ikx} dk \\ &= \frac{1}{\sqrt{2\pi D(t-t_0)}} \exp\left[-\frac{x^2}{4D(t-t_0)}\right] \end{aligned} \tag{5.30}$$

where the prefactor comes from satisfying normalisation, so the area under the curve (surface) is 1. Moreover, $n = 0$ when $t < t_0$. ☐

Again we have shown that a random selection of particles obey some form of distribution law, namely the distribution of particles forms a Gaussian.

Corollary 5.2.1. *We recover the definition of the Dirac-δ function by taking limits. Since $t \geq t_0$ we take limits $t \to t_0^+$ so*

$$\lim_{t \downarrow t_0} n(x,t) = \delta(t-t_0)\delta(x-x_0)$$

Corollary 5.2.2. *With an advective term, the fundamental solution is*

$$n(x,t) = \frac{1}{\sqrt{2\pi D(t-t_0)}} \exp\left[-\frac{(x-vt)^2}{4D(t-t_0)}\right]$$

Statistically, this is a normal distribution with mean $\overline{x} = vt$ and standard deviation $\sigma = 2D(t-t_0)$. Furthermore, the Full-Width-Half-Maximum (FWHM) is 2σ wide.

Theorem 5.2.4. *The solution to the same problem in 3D (i.e. the 3D diffusion equation) for particles centred at some $\vec{x_0}$ is*

$$n(\vec{x},t) = (2\pi D(t-t_0))^{-3/2} \exp\left[-\frac{|\vec{x}-\vec{x}_0-\vec{v}(t-t_0)|^2}{4D(t-t_0)}\right] \tag{5.31}$$

Proof. Due to properties of Fourier transforms, the different directions are independent - namely our solution is separable. ☐

Physically, this means each direction is independent, and does not affect distributions in the other directions. In this case, $\sigma \sim \sqrt{Dt}$ and

$$\langle |\vec{x} - \vec{x_0}^2| \rangle \sim Dt \tag{5.32}$$

Moreover, suppose we continue to let the system evolve and make many measurements. Then the error, assuming they are not systematic, i.e. they are random, will keep growing. Importantly,

$$\text{Error} \sim \sqrt{N} \implies \frac{\text{Error}}{N} \sim \frac{1}{\sqrt{N}} \tag{5.33}$$

Generally speaking, If you find an argument to show that the time variation of some quantity is **stochastic** (governed by random forces or charges) characterised by a size of change a and a rate of events $1/\tau$, then you can safely assume the solution is a Gaussian for n as in Eq. (5.31) with $D \sim a^2/\tau$.

However there are caveats to take into account:

- Is D a constant? It is common for D to depend on x, it may not change how the system behaves after a long time but it can make things complicated (e.g., we can't pull D out in the continuity equation).
- Are jumps large? Need to think about the distribution in a like in financial markets. It is helpful to carefully consider what is being modelled.

Finally, consider the diffusion equation

$$\frac{\partial n}{\partial t} - D\frac{\partial^2 n}{\partial x^2} = \delta(t-t_0)\delta(x-x_0) \tag{5.34}$$

which is the diffusion equation with an initial particle at (x_0, t_0) (which is what the RHS specifies physically). This is a linear PDE and that is awesome, but we could have a more complicated inhomogeneous PDE

$$\frac{\partial n}{\partial t} - D\frac{\partial^2 n}{\partial x^2} = \delta(t-t_0)f(x) \tag{5.35}$$

which is saying at $t = t_0$ we have some distribution of particles $f(x)$. The fact it is linear means we can treat $f(x)$ as an integral over the sum of many weighted δ functions at different x values - namely we must do an integral and

$$n(x,t) = \int G(x-x_0, t-t_0)f(x_0)dx_0 \tag{5.36}$$

where

$$G \sim \frac{1}{\sqrt{2\pi D(t-t_0)}}\exp\left[-\frac{(x-x_0)^2}{4D(t-t_0)}\right]$$

We could also have some function of t so now we'd also need to integrate over time, but over the interval $[-\infty, t]$ since we cannot have future events influence present events.

5.2.2 Momentum Diffusion

This section is **NON-EXAMINABLE FOR 2023-24** and therefore has currently been omitted.

5.3 Black-Scholes Equation

This is an application of diffusion and drift.

Definition 5.3.1. A **Call Option** is the *right* (but no obligation) to buy the share (or whatever financial asset) at a fixed **strike price** K on a fixed future **maturity** time T_m.

Definition 5.3.2. A **put** option is the right (not obligation) to sell the share at some strike price S_m and maturity time T_m

Suppose I'm selling an NVIDIA RTX 4090 (a GPU). Suppose I fix the price at £1700 based on current markets and you request from me a call option, say a couple months you want it to be £1600. If the market price drifts below £1600 towards that maturity date, you are less likely to purchase from me at your chosen time. We want to be able to model asset prices.

The argument for the model goes as so

1. Consider logarithm $x = \log(S)$ (called the **logprice**) of the share or asset price S, equivalent to fractional changes in S. Namely, $\delta x = \delta S/S$..
2. We assume an **efficient market**: any anticipatable coming price change should already be priced in, so the expected price change should simply match general interest rates (also called **discounting**).
3. Every time interval $\delta t = \sum_i \delta t_i$ with corresponding changes in logprice $\delta x = \sum_i \delta x_i = \delta S/S$, the δx_i should be **statistically independent** leading to δx being **Gaussian distributed** (by the Central Limit Theorem). This assumes the δx_i have finite variances.

The share price function

$$\langle S(t)\rangle = S(t_0)e^{r(t-t_0)} = S(t_0)\exp\left[\int_{t_0}^{t} r(t')dt'\right] \tag{5.37}$$

where r is the **rate of interest**. The fractional **expected price update** in a small interval

$$\langle S(t+\delta t)/S(t)\rangle = \langle e^{\delta x}\rangle = e^{D\delta t + v\delta t} \tag{5.38}$$

The probability of logprice change

$$P(\delta x, \delta t) \propto \exp\left(-\frac{(\delta x - v\delta t)^2}{4D\delta t}\right)$$

where v is drift velocity of price. This is once again a Gaussian distribution with

Definition 5.3.3. The **volatility** is defined as

$$2D = \frac{\text{variance of logprice change}}{\delta t} \tag{5.39}$$

The volatility is commonly *cited* as

$$\frac{\sigma}{\sqrt{\delta t}} = \sqrt{2D}$$

Comparing Eqs. (5.37),(5.38) we see that

$$v = -D + r \tag{5.40}$$

Notice our v has a $-D$ term, which is due to working with logprice rather than actual price.

Stocks themselves are *conserved*. 1 share in a company remains, 1 share in a company. It is possible for a company to increase the number of shares but this generally has to be agreed on by all investors in the company. But if there is no motivation to do so (e.g., no 'force'), then shares are conserved even if their prices changes. This conservation of stocks is like conservation of particles - one electron remains 1 electron. We make the analogue

$$\text{Position of particle} \longleftrightarrow \text{Logprice of stock/asset} \tag{5.41}$$

Now, we identify **probability current density**

$$j(x,t) = vP - D\frac{\partial P}{\partial x}$$

Substituting into the continuity equation gives us another Fokker-Planck equation:

$$\frac{\partial P(x,t|x_0,t_0)}{\partial t} + \frac{\partial}{\partial x}\left[(-D+r)P\right] = \frac{\partial}{\partial x}D\frac{\partial P}{\partial x} \tag{5.42}$$

with the initial condition $P(x,t_0) = \delta(x-x_0)$ that at present time t_0 we have the known log price x_0

Definition 5.3.4. $V_m(S)$ is the **payoff at a maturity time** T_m:

$$V_m(S) = \begin{cases} 0 & S < K \\ (S-K) & S \geq K \end{cases}$$

Namely, we know the strike price K, and to satisfy our call option, we want S to increase if we are selling. Call options on materials can also be thought of in terms of insurance: avoids paying higher than K.

- If $S_m > K$ you will exercise your call option to *buy* (because you can make a return later)
- if $S_m < K$ your counterparty in the put option will exercise their right to *sell* (since they can now make a return)

Definition 5.3.5. Suppose at some later time $t \leq t_m$ we have established the fair value to be $V(x,t) = V_m(S = e^x)$, i.e. it depends on the logprice x at that time. This is the **future condition**.

Averaging over those future values conditional on price at time t_0, the fair value at time t_0 is then

$$V\left(x_0, t_0\right) = e^{-\int_{t_0}^{t} r(t')dt'} \int P\left(x, t \mid x_0, t_0\right) V(x,t) dx$$

Provided $t_0 < t < t_m$ our valuation $V(x,t)$ should not depend on the choice of intermediate time t, so ∂_t of the RHS should be zero. This leads to

$$0 = \partial_t V(x_0, t_0) = e^{-r(t-t_0)} \int -rPV + \partial_t P + V\partial_t P dx$$

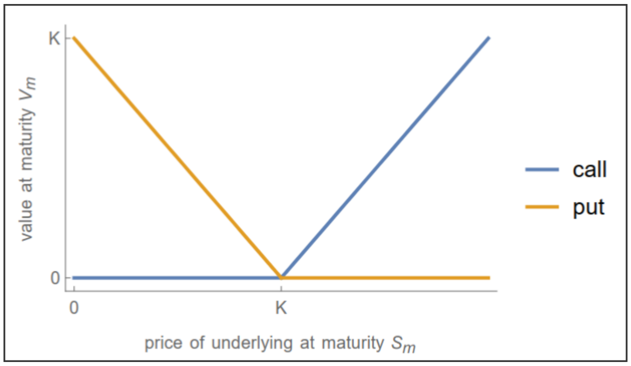

Figure 5.1: The value of simple options vs price S at maturity. The holder of a call (put) option has the right to buy (sell) at the "strike" price K on the maturity date

Integrating by parts (twice) with respect to x then leads to

$$0 = \int P(x - x_0, t - t_0) \left(\mathcal{D}^A V + \frac{\partial V}{\partial t} - rV \right) dx$$

where the **adjoint operator** is given by $\mathcal{D}P$ with $\mathcal{D}^A = \partial_x D \partial_x + \partial_x(-D + r)$.

The fair valuation $V(x, t)$ depends only on the future evolution from t to maturity whereas the factor P independently varies with conditions set at time t_0. Hence we need

$$\frac{\partial V}{\partial t} - \mathcal{D}^A V - rV = 0 \tag{5.43}$$

for the time evolution of the option value, propagating backwards from its value $V_m(x)$ at maturity. Substituting $\mathcal{D}^A$ gives

Black-Scholes equation

$$\frac{\partial V}{\partial t} + S^2 \frac{\partial}{\partial S} D \frac{\partial}{\partial S} V + r \left(S \frac{\partial}{\partial S} - 1 \right) V = 0 \tag{5.44}$$

The sign of the $\frac{\partial V}{\partial t}$ term in the Black-Scholes equation is important. It only makes sense as a diffusion equation when integrating backwards in time.

Definition 5.3.6. We define $\tau = t_m - t$ as the **time to maturity** where t_m is the time *of* maturity.

Therefore the Black-Scholes equation is a Fokker-Planck equation with the initial condition $V(x, \tau = 0) = V_m(S = e^x)$. A graph of V against S like in Fig. 5.1 now smooths out the corner with increasing τ.

Theorem 5.3.1. *An explicit general solution of the Black-Scholes equation is of the form*

$$V(x, \tau) = \int G(x - x_m, \tau) V_m(e^{x_m}, \tau = 0) dx_m \tag{5.45}$$

where G the Green function obeys

$$\frac{\partial G}{\partial \tau} = D\frac{\partial^2 G}{\partial x^2} - (D-r)\frac{\partial G}{\partial x} - rG + \delta(\tau)\delta(x - x_m) \tag{5.46}$$

with

$$G = \frac{1}{\sqrt{4\pi D\tau}} \exp\left[-\frac{(x-(x_m+(D-r)\tau))^2}{4D\tau}\right] \exp\left[-r\tau\right] \tag{5.47}$$

where the second exponential comes from the fact that the further away we go back in time from maturity, the more disappointed we are in our returns without interest (i.e. discounting).

The drift term $-(D-r)$ is time reversed otherwise the exponential should have a minus sign not a plus sign.

G tells you the contribution to the value of the option from the scenario where the logprice is x_m at maturity.

We can have more exotic options:

Call option with floor If $S < S_{\text{floor}}$ option is automatically void. In Physics, this would be implemented as a boundary condition like $V = 0$ at $S_{\text{floor}} = e^x$. This is set so the owner of the share doesn't make too huge of a loss selling the share to you.

Put option with barrier If $S > S_{\text{barrier}}$ option is automatically void. In Physics, this would be implemented as a boundary condition like $V = 0$ at $S_{\text{barrier}} < e^x$. This is set so a share price can't get so large, otherwise no one would buy.

If you have a call option with a barrier at high price that kills the contract when you have a large value, you don't strictly want that because of this cap. However it makes the share easier to sell as it is less appealing, so it is important to weigh options.

In an American/Asian option, we can exercise ANYTIME up to maturity. Exercising early can be favourable so that there is less discounting. For the strike price, in a call option, you pay this, so it is good to *delay*. For a put option, shares are received so it is nice to get a payoff earlier.

One of the big questions is *when should you exercise*?. If S is far from K very early on, you should exercise as soon as possible the further away it is, i.e. there is some locus which decides when you should exercise your options. However, there is no closed form to this locus.

Chapter 6

Kinetic Theory of Plasma

This section is **non-examinable for 2023-24** and will be printed here at a later date.

Chapter 7

Credits

This was produced on behalf of Warwick Physics Society. Contributors:

- Chung Xu (Main editor)
- Sam Deacon