

CONDENSED MATTER THEORY

PETER DANIEL JOHANNSSEN



Summary of my Studies at the University of Copenhagen

STATISTICAL PHYSICS

I.1 MULTIPLICITY, ENTROPY ETC.

Multiplicity is defined in Kittel as the number of (quantum) states with very nearly the same energy. However, other authors such as Schroeder prefer to define it as the number of microstates that correspond to any given macrostate.

Binary systems: An example of a system that is simple enough for us to write an expression for the multiplicity are binary systems, for example 1/2-spin particles, whose spin either points upwards (\uparrow) or downwards (\downarrow). In this case the multiplicity, $g(N, N_\uparrow, N_\downarrow)$, is given by

$$g(N, N_\uparrow, N_\downarrow) = \frac{N!}{N_\uparrow! N_\downarrow!} = \frac{N!}{\left(\frac{1}{2}N + s\right)! \left(\frac{1}{2}N - s\right)!} \quad (1.1)$$

where $2s \equiv N_\uparrow - N_\downarrow$. Using **Stirling's Approximation** this allows to approximate the multiplicity of binary systems as a gaussian distribution *for large N*:

$$g(N, s) \approx g(N, 0) \exp\left(-\frac{2s^2}{N}\right), \quad (1.2)$$

$$g(N, 0) \approx \left(\frac{2}{\pi N}\right)^{\frac{1}{2}} 2^N \quad (1.3)$$

The energy of a **binary spin system** consisting of N spin-1/2 particles is

$$U = - \sum_{i=1}^N \mathbf{m}_i \cdot \mathbf{B} = -mB(N_\uparrow - N_\downarrow) = -2msB \quad (1.4)$$

where \mathbf{m} is the single magnetic moment, and \mathbf{B} the applied magnetic field.. We can also define the total magnetic moment $M = 2ms$

Multiplicity function for N harmonic oscillators: (Einstein solid). Suppose we have n energy quanti, thus that the system has total energy $\varepsilon = n\hbar\omega$, then the multiplicity is given as

$$g(N, n) = \frac{(N + n - 1)}{n!(N - 1)!} \quad (1.5)$$

Averages Let's suppose a physical property, X , has the value $X(s)$ when in the state s , then the average value of X will be:

$$\langle X \rangle = \sum_s X(s)P(s), \quad \sum_s P(s) = 1 \quad (1.6)$$

where $P(s)$ is the probability of finding the system in state s .

Interacting systems:

Given two systems with N_1 and N_2 particles respectively. If their respective multiplicities are given by $g_1(N_1, s_1)$ and $g_2(N_2, s_2)$ and we hold $s = s_1 + s_2$ constant, then the systems combined multiplicity is

$$g(N, s) = \sum_{s_1} g_1(N_1, s_1) g_2(N_2, s - s_1) \quad (1.7)$$

Two spin systems in thermal contact: Once again using the Stirling Approximation we can approximate the multiplicity of two interacting spin-1/2 systems as gaussian in the variables s_1 and s_2 :

$$g_1(N_1, s_1) g_2(N_2, s_2) = g_1(0) g_2(0) \exp\left(-\frac{2s_1^2}{N_1} - \frac{2s_2^2}{N_2}\right) \quad (1.8)$$

$$= g_1(0) g_2(0) \exp\left(-\frac{2s_1^2}{N_1} - \frac{2(s - s_1)^2}{N_2}\right) \quad (1.9)$$

where $g_i(0)$ are given by the same expression we had before for independent binary systems. By taking the first and second derivative of this with respect to s_1 we see that the maximum (and thus the most likely state) is achieved when

$$\frac{s_1}{N_1} = \frac{s_2}{N_2} \quad (1.10)$$

Thermal equilibrium and temperature: If we more generally write g as a function of N and U we will see that at thermal equilibrium (i.e state of maximal multiplicity) the following equation holds:

$$\frac{1}{g_1} \left(\frac{\partial g_1}{\partial U_1} \right)_{N_1} = \frac{1}{g_2} \left(\frac{\partial g_2}{\partial U_2} \right)_{N_2} \quad (1.11)$$

here we see that it is natural to introduce $\ln g(N, U) = \sigma$, because that transforms the equation above into:

$$\left(\frac{\partial \sigma_1}{\partial U_1} \right)_{N_1} = \left(\frac{\partial \sigma_2}{\partial U_2} \right)_{N_2} \quad (1.12)$$

This leads naturally to the concept of temperature; we define temperature as

$$\frac{1}{\tau} = \left(\frac{\partial \sigma}{\partial U} \right)_N, \quad \tau = k_B T \quad (1.13)$$

which agrees with the fact that two systems in thermal equilibrium will have the same temperature. σ is the unitless entropy, and is related to the entropy in Schroeder by the simple relation

$$S = k_B \sigma = k_B \ln g \quad (1.14)$$

Laws of thermodynamics:

0. If two systems are in thermal equilibrium with a third system, then they must be in thermal equilibrium with each other.
1. Heat is a form of energy
2. If a closed system is in a configuration that is not the equilibrium configuration, the most probable consequence will be that the entropy of the system will increase monotonically in successive instants of time.
3. $\lim_{T \rightarrow 0} S(T) = 0$.

I.2 BOLTZMANN DISTRIBUTION & HELMHOLTZ FREE ENERGY

Partition function:

$$Z \equiv \sum_s \exp\left(-\frac{\epsilon_s}{\tau}\right) \quad (1.15)$$

where each individual term is referred to as a Boltzmann factor. The **probability** to find a system is state s :

$$P(\epsilon_s) = \frac{\exp(-\epsilon_s/\tau)}{Z} \quad (1.16)$$

from here we see that $\sum_s P(\epsilon_s) = 1$, as expected. With the partition function we can calculate many useful quantities, such as the total energy of the system:

$$U = \tau^2 \frac{\partial \ln Z}{\partial \tau} \quad (1.17)$$

Legendre Transformation

$$dU = \tau d\sigma - p dV \quad (1.18)$$

from here we see that U is naturally a function of σ and V . However, using a Legendre Transformation we can transform U into a new function F which is a function of τ and V :

$$F \equiv U - \tau\sigma \rightsquigarrow dF = -\sigma d\tau - p dV \quad (1.19)$$

F is the **Helmholtz Free Energy** and can be more useful than U , for instance in cases where we know τ and not σ . F is *minimized* at equilibrium for constant volume (because σ is *maximised*).

From our expression for dF we see that

$$\left(\frac{\partial F}{\partial \tau}\right)_V = -\sigma, \quad \left(\frac{\partial F}{\partial V}\right)_\tau = -p \quad (1.20)$$

Another useful property of F is that it is closely related to the partition function:

$$F = -\tau \ln Z \quad (1.21)$$

Ideal Gas: An ideal gas is a gas of non-interacting particles confined in a box. Thus each particle has energy

$$\epsilon_n = \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2), \quad n^2 = n_x^2 + n_y^2 + n_z^2 \quad (1.22)$$

In which case

$$Z = \sum_{n_x} \sum_{n_y} \sum_{n_z} \exp\left(-\frac{\hbar^2 \pi^2 (n_x^2 + n_y^2 + n_z^2)}{2mL^2 \tau}\right) \quad (1.23)$$

which can be approximated with three integrals, which gives

$$Z = \frac{\pi^{\frac{3}{2}}}{8} \left(\frac{2mL^2 \tau}{\hbar^2 \pi^2}\right)^{\frac{3}{2}} = \frac{V}{(2\pi \hbar^2 / m \tau)^{\frac{3}{2}}} = n_Q V \quad (1.24)$$

where we've defined the **quantum concentration**:

$$n_Q = \left(\frac{m\tau}{2\pi\hbar^2} \right)^{\frac{3}{2}} \quad (1.25)$$

This is the concentration associated with an atom in a cube with side length equal to the thermal average de Broglie length. From here we can define

Classical Regime: $n \ll n_Q$

where $n = V^{-1}$.

For a **system of N indistinguishable particles** the combined partition function is:

$$Z_N = \frac{Z_1^N}{N!} \quad (1.26)$$

where the $N!$ accounts for us being able to interchange particles without affecting the partition function. This interchangeability has been shown to be true experimentally.

From the partition function of N indistinguishable particles we can calculate the energy and entropy of an ideal gas:

$$U = \tau^2 \left(\frac{\partial \ln Z_N}{\partial \tau} \right) = \frac{3}{2} N \tau \quad (1.27)$$

which is just the kinetic part of the equipartition theorem. Additionally using that $F = -\tau \ln Z$ and that $\sigma = -\left(\frac{\partial F}{\partial \tau}\right)_V$ we get the **Sackur-Tetrode Equation**:

$$\sigma = N \left(\ln \left(\frac{n_Q}{n} \right) + \frac{5}{2} \right) \quad (1.28)$$

Entropy of mixing: A system with N particles of type A and $N - t$ particles of type B has the following multiplicity:

$$g(N, t) = \frac{N!}{(N - t)!t!} \quad (1.29)$$

as seen previously. The entropy associated with this system is therefore given by

$$\sigma = \ln N! - \ln(N - t)! - \ln t! \quad (1.30)$$

$$\approx -(N - t) \ln(1 - t/N) - t \ln t/N \quad (1.31)$$

where I have used Stirlings Approximation. Letting $x = t/N$ we get

$$\sigma(x) = -N ((1 - x) \ln(1 - x) - x \ln x) \quad (1.32)$$

which looks a lot like a concave parabola, reaching its maximum at $x = 1/2$, i.e. there where there are exactly as many particles of species A as there are of species B.

1.3 THERMAL RADIATION

Planck distribution function: Let ε_s denote the energy that is associated with s photons each with (angular) frequency ω , that is

$$\varepsilon_s = s\hbar\omega \quad (1.33)$$

The partition function for this system is

$$Z = \sum_{s=0}^{\infty} \exp(-\hbar\omega s/\tau) = \frac{1}{1 - \exp(-\hbar\omega/\tau)} \quad (1.34)$$

Thus the probability to find the system in state s is

$$P(s) = \frac{\exp(-\hbar\omega s/\tau)}{Z} \quad (1.35)$$

The thermal average of s is

$$\langle s \rangle = \sum_{s=0}^{\infty} sP(s) = Z^{-1} \sum_{s=0}^{\infty} s \exp(-\hbar\omega s/\tau) \quad (1.36)$$

$$= \frac{1}{\exp(\hbar\omega/\tau) - 1} \quad (1.37)$$

which is known as the **Planck distribution function**. Using this distribution function we can calculate the total energy of the system:

$$U = \sum_{n_x} \sum_{n_y} \sum_{n_z} \frac{\hbar\omega_n}{\exp(\hbar\omega_n) - 1} \approx \frac{V\pi^2}{15\hbar^3 c^3} \tau^4 \quad (1.38)$$

where I have used a method that is described in the appendix to approximate the triple sum as an integral. The result,

$$\frac{U}{V} = \frac{\pi^2}{15\hbar^3 c^3} \tau^4 \quad (1.39)$$

is known as the **Stefan-Boltzmann law of radiation**. Additionally the integrand from the expression that is used to calculate the Stefan-Boltzmann law is the spectral density and is given as

$$u_{\omega} = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\exp(\hbar\omega/\tau) - 1} \quad (1.40)$$

which is the well known distribution that describes radiation from black bodies, as can be seen in the following figure

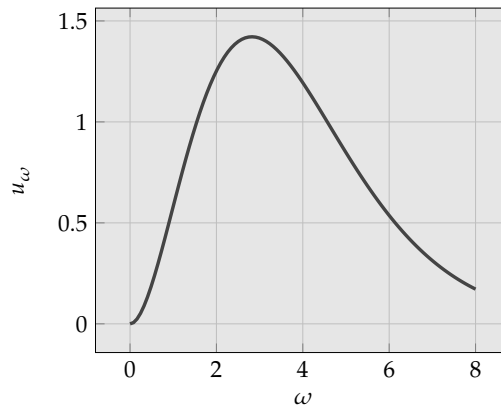


Figure 1.1: Planck Distribution Function

Phonons: For phonons there is a limit to the total number of modes: $3N$. Therefore

$$\frac{3}{8} \int_0^{n_D} 4\pi n^2 dn = 3N \quad (1.41)$$

where the 3 on the left side of the equality is because there phonons have three degrees of freedom (they can propagate in three directions in space). From there we get

$$n_D = \left(\frac{6N}{\pi} \right)^{\frac{1}{3}} \quad (1.42)$$

and the total energy of this system is

$$U = \frac{3\pi}{2} \int_0^{n_D} dn \, n^2 \frac{\hbar\omega_n}{\exp(\hbar\omega_n/\tau) - 1} \quad (1.43)$$

which in the low temperature limit becomes

$$U \approx \frac{3\pi^4 N k_B T^4}{5\theta^3}, \quad \theta = \frac{\hbar v}{k_B} \left(\frac{6\pi^2 N}{V} \right)^{\frac{1}{3}} \quad (1.44)$$

where v is the velocity of sound in the medium. This thus leads us to an important result

$$C_V = \left(\frac{\partial U}{\partial \tau} \right)_V = \frac{12\pi^4 N}{5} \left(\frac{\tau}{k_B \theta} \right)^3, \quad \tau \ll k_B \theta \quad (1.45)$$

The τ^3 dependancy was predicted experimentally, but couldn't be explained theoretically prior to **Debye**.

1.4 CHEMICAL POTENTIAL

Let's imagine two systems \mathcal{S}_1 and \mathcal{S}_2 that are in thermal contact with each other, and with a reservoir. Additionally the systems can interchange particles. If the systems are in thermal equilibrium and their volumes are constant then

$$\left(\frac{\partial F_1}{\partial N_1} \right)_\tau = \left(\frac{\partial F_2}{\partial N_2} \right)_\tau \quad (1.46)$$

at chemical equilibrium. From here we define the **chemical potential**:

$$\mu \equiv \left(\frac{\partial F}{\partial N} \right)_{\tau, V} \quad (1.47)$$

thus $\mu_1 = \mu_2$ at chemical equilibrium. Note that particles will tend to "flow" from high chemical potential to low chemical potential. (Just like energy flows from high temperature to low temperature).

For an ideal gas

$$\mu = \tau \ln \left(\frac{n}{n_Q} \right) \quad (1.48)$$

which can be found directly from our prior definition of Z for an ideal gas.

Internal and external chemical potential: External potentials will offset the chemical potential of a system:

$$\mu = \mu_{int} + \mu_{ext} \quad (1.49)$$

For instance **an ideal gas in a gravitational potential**:

$$\mu = \tau \ln \left(\frac{n}{n_Q} \right) + mgh \quad (1.50)$$

here we can already see that particles will tend to "flow" downwards, there where the gravitational potential is lowest. In equilibrium the chemical potential is independent of height, and thus we can get that

$$n(h) = n(0) \exp(-mgh/\tau) \quad (1.51)$$

which of course only is true close to the surface of earth.

Another example is **magnetic dipoles in a uniform magnetic field**:

$$\mu_{tot}(\uparrow) = \tau \ln \left(\frac{n_{\uparrow}}{n_Q} \right) - mB \quad (1.52)$$

$$\mu_{tot}(\downarrow) = \tau \ln \left(\frac{n_{\downarrow}}{n_Q} \right) + mB \quad (1.53)$$

Again using our equilibrium condition that $\mu_{tot}(\uparrow) = \mu_{tot}(\downarrow)$ we get that

$$n(B) = n(0) \cosh(mB/\tau) \quad (1.54)$$

concentration increases with increasing magnetic field, which you can observe by putting iron dust close to a bar magnet.

Extending the thermodynamic identity:

$$d\sigma = \left(\frac{\partial \sigma}{\partial U} \right)_{V,N} dU + \left(\frac{\partial \sigma}{\partial V} \right)_{U,N} dV + \left(\frac{\partial \sigma}{\partial N} \right)_{V,U} dN \quad (1.55)$$

which can be written as

$$d\sigma = \frac{1}{\tau} dU + \frac{p}{\tau} dV - \frac{\mu}{\tau} dN \quad (1.56)$$

or, equivalently

$$dU = \tau d\sigma - p dV + \mu dN \quad (1.57)$$

Gibbs Factor and Gibbs Sum:

A system that is in thermal and chemical contact with a reservoir can be described using the Gibbs sum (that is to say that the system can interchange *energy* and *particles* with a reservoir). The **Gibbs sum** or **Grand sum** is defined as

$$\mathcal{Z} = \sum_{N=0}^{\infty} \sum_{s(N)} \exp \left(\frac{N\mu - \varepsilon_{s(N)}}{\tau} \right) \quad (1.58)$$

Which, like with the partition function, allows us to define the probability of finding the system in state N_1, ε_1 :

$$P(N_1, \varepsilon_1) = \frac{\exp(N_1\mu - \varepsilon_1)}{\mathcal{Z}} \quad (1.59)$$

Once again it is clear that $\sum_N \sum_{s(N)} P(N, s(N)) = 1$ Often one defines the **absolute activity**:

$$\lambda = \exp(\mu/\tau) \quad (1.60)$$

The average number of particles is thus

$$\langle N \rangle = \lambda \frac{\partial}{\partial \lambda} \ln Z = \tau \frac{\partial \ln Z}{\partial \mu} \quad (1.61)$$

and the total energy is

$$U = \langle \epsilon \rangle = \left(\tau \mu \frac{\partial}{\partial \mu} + \tau^2 \frac{\partial}{\partial \tau} \right) \ln Z \quad (1.62)$$

An example where Kittel uses the Gibbs sum can be seen on page 141-144.

Often we use $f(\epsilon)$ interchangeably with $\langle N \rangle$:

$$f(\epsilon) = \langle N \rangle \quad (1.63)$$

Fermi-Dirac distribution:

$$f(\epsilon) = \frac{1}{\exp((\epsilon - \mu)/\tau) + 1} \quad (1.64)$$

Used to describe *fermions* – particles with half integer spin, that are affected the Pauli principle. We define the Fermi energy as

$$\epsilon_F = \mu(\tau = 0) \quad (1.65)$$

Note that

$$\lim_{\tau \rightarrow 0} \frac{1}{\exp((\epsilon - \mu)/\tau) + 1} = \begin{cases} 1 & \text{for } \epsilon < \epsilon_F \\ 0 & \text{for } \epsilon > \epsilon_F \end{cases} \quad (1.66)$$

Bose-Einstein Distribution

For particles with integer spin (bosons) the Pauli exclusion principle does not hold, which means that different particles *can* occupy the same state. This implies that as $\tau \rightarrow 0$ more and more particles will be in the ground state (at $\tau = 0$ they will *all* be in the ground state). For bosons we use:

$$f(\epsilon) = \frac{1}{\exp((\epsilon - \mu)/\tau) - 1} \quad (1.67)$$

Classical limit:

Note that both distribution functions behave identically in the classical limit. The classical limit is defined as

$$\exp\left(\frac{\epsilon - \mu}{\tau}\right) \gg 1 \quad (1.68)$$

Here both distributions are approximately

$$f(\epsilon) \approx \lambda \exp\left(-\frac{\epsilon}{\tau}\right) \quad (1.69)$$

known as the **classical distribution function**.

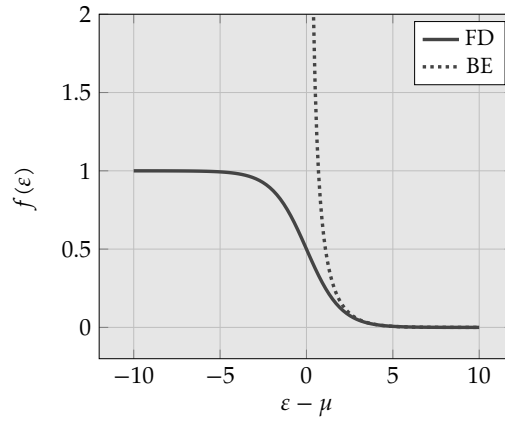


Figure 1.2: Fermi-Dirac distribution (dotted) and the Bose-Einstein distribution

Ideal Gas:

We previously related Z of an ideal gas to the concentration, from here we can define the chemical potential of an ideal gas as

$$\mu = \tau \ln \left(\frac{n}{n_Q} \right) \quad (1.70)$$

Internal degrees of freedom

If the energy of an ideal gas can be written as

$$\epsilon = \epsilon_n + \epsilon_{int} \quad (1.71)$$

where ϵ_n is the old energy and ϵ_{int} is some additional internal energy (perhaps due to rotational kinetic energy), then we can make the substitution $\epsilon_n \rightsquigarrow \epsilon_n + \epsilon_{int}$ in all our prior results. This would modify our Gibbs sum to

$$Z = 1 + \lambda Z_{int} \exp(-\epsilon_n/\tau), \quad (1.72)$$

$$Z_{int} = \sum_{int} \exp(-\epsilon_{int}/\tau) \quad (1.73)$$

For more on this see page 170-171.

Expansion of ideal gases

	ΔU	$\Delta \sigma$	δW	δQ
RITE	0	$N \ln \frac{V_2}{V_1}$	$-N\tau \ln \frac{V_2}{V_1}$	$N\tau \ln \frac{V_2}{V_1}$
RISE	$-\frac{3}{2}N\tau_1 \left(1 - \left(\frac{V_1}{V_2} \right)^{\frac{2}{3}} \right)$	0	$-\frac{3}{2}N\tau_1 \left(1 - \left(\frac{V_1}{V_2} \right)^{\frac{2}{3}} \right)$	0
IE	0	$N \ln \frac{N_2}{N_1}$	0	0

Table 1.1: RITE: Reversible isothermal expansion, RISE: reversible isentropic expansion, IE: irreversible expansion

1.5 FERMI & BOSE GASES

Quantum gas

A quantum gas is defined as a gas in the limit

$$n \geq n_Q \equiv \left(\frac{m\tau}{2\pi\hbar^2} \right)^{\frac{3}{2}} \quad (1.74)$$

Or, if we have fixed the concentration, then a gas becomes quantum (or degenerate) when

$$\tau < \tau_Q = \frac{2\pi\hbar^2}{m} n^{\frac{2}{3}} \quad (1.75)$$

Ground state of Fermi gas (i.e $\tau = 0$) The highest energy level is

$$\epsilon_F = \frac{\hbar^2}{2m} \left(\frac{\pi n_F}{L} \right)^2 = \frac{\hbar^2}{2m} \left(\frac{3\pi^2 N}{V} \right)^{\frac{2}{3}}, \quad (1.76)$$

$$n_F = \left(\frac{3N}{\pi} \right)^{\frac{1}{3}} \quad (1.77)$$

The total energy

$$U_0 = 2 \sum_n \epsilon_n = \frac{\pi^3}{2m} \left(\frac{\hbar}{L} \right)^2 \int_0^{n_F} dn n^4 = \frac{3}{5} N \epsilon_F \quad (1.78)$$

Don't confuse n here for N/V . **Density of states** Instead converting the sum into an integral over n it can be useful to integrate over ϵ . However, the integrand then needs to be weighted according to how many particles have the energy ϵ :

$$\sum_n (\dots) \rightsquigarrow \int d\epsilon \mathcal{D}(\epsilon) (\dots) \quad (1.79)$$

where $\mathcal{D}(\epsilon)$ is the **density of states**. We can calculate it using

$$\mathcal{D}(\epsilon) \equiv \frac{dN}{d\epsilon} \quad (1.80)$$

For instance a 3D fermi gas has the following density of states:

$$\mathcal{D}(\epsilon) = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar} \right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}} \quad (1.81)$$

For 1D and 2D expressions use the length of a line or area of a circle instead of the surface of a sphere when deriving equation 7.6 in the book.

Heat Capacity of electron gas

The energy increase of an electron gas when heated from 0 to τ can be calculated using $\mathcal{D}(\epsilon)$:

$$\Delta U = \int_{\epsilon_F}^{\infty} d\epsilon \epsilon \mathcal{D}(\epsilon) f(\epsilon) \quad (1.82)$$

and from there we can get the following expression for the heat capacity of an electron gas:

$$C_{el} = \begin{cases} \frac{1}{3}\pi^2 \mathcal{D}(\epsilon_F) \tau & \text{for } \tau \ll \tau_F \\ \frac{1}{2}\pi^2 N \frac{\tau}{\tau_F} & \text{for } \tau \gg \tau_F \end{cases} \quad (1.83)$$

Heat Capacity of metals

$$C_V = \gamma \tau + A \tau^3 \quad (1.84)$$

where γ and A are constants characteristic of the material. Note that the linear term is due to the electrons whereas the cubic term is due to lattice vibrations.

For a short note on white dwarf stars and nuclear matter see pages 196-199.

Boson gas an Einstein condensation

When a boson gas reaches very low temperatures a substantial fraction of the total number of particles will occupy the ground state – this phenomenon is known as **Einstein condensation**. The chemical potential for an Einstein condensate is

$$\mu = -\frac{\tau}{N} \quad (1.85)$$

The chemical potential of a boson system must always be lower in energy than the ground state energy in order for the occupancy of every orbital to be non-negative.

Occupancy of the ground state of a boson gas

We can split the particles into particles that are in the ground state, of which there are N_0 , and particles that are excited, of which there are N_e :

$$N = N_0(\tau) + N_e(\tau) = N_0(\tau) + \int_0^\infty d\epsilon \mathcal{D}(\epsilon) f(\epsilon, \tau) \quad (1.86)$$

or

$$N_e = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \tau^{\frac{3}{2}} \int_0^\infty dx \frac{x^{\frac{1}{2}}}{\lambda^{-1} e^x - 1} \quad (1.87)$$

from where we get that

$$\frac{N_e}{N} \approx 2.612 \frac{n_Q}{n} \quad (1.88)$$

The temperature where $N = N_e$ is the **Einstein condensation temperature**:

$$\tau_E \equiv \frac{2\pi\hbar^2}{m} \left(\frac{N}{2.612V} \right)^{\frac{2}{3}} \quad (1.89)$$

This might seem like an odd definition, but as soon as $\tau < \tau_E$ there number of particles in the ground state becomes significant. This allows us to write

$$\frac{N_e}{N} \approx \left(\frac{\tau}{\tau_E} \right)^{\frac{3}{2}} \quad (1.90)$$

1.6 GIBBS FREE ENERGY

The Gibbs free energy is defined as

$$G \equiv U - \tau\sigma + pV \quad (1.91)$$

Note that this is a (double) Legendre transformation, where we have replaced the variables σ and V (which U are a function of) with τ and P , thus

$$dG = -\sigma d\tau + V dp + \mu dN \quad (1.92)$$

which implies that

$$\left(\frac{\partial G}{\partial \tau}\right)_{p,N} = -\sigma, \quad \left(\frac{\partial G}{\partial p}\right)_{\tau,N} = V \quad (1.93)$$

$$\left(\frac{\partial G}{\partial N}\right)_{\tau,p} = \mu \quad (1.94)$$

Due to the fact that G is a function of only N, p and τ and p, τ are *intensive quantities* (see Appendix for a definition of these) we know that G has to be linearly proportional to N , and using the partial derivative above we get that

$$G = N\mu(\tau, p) \quad (1.95)$$

Equilibrium in (chemical) reactions

Suppose we write a chemical reaction as

$$\sum_i v_i A_i = 0 \quad (1.96)$$

where A_i is just the name of the chemical (this means products have *negative* v_i and reactants have *positive* v_i). Due to the fact that $dG = 0$ at equilibrium we infer that

$$\sum_i \mu_i v_i = 0 \quad (1.97)$$

Ideal gasses For ideal gasses we have that

$$\mu_j = \tau(\ln n_j + \ln c_j), \quad c_j \equiv n_{Qj} Z_j^{\text{int}} \quad (1.98)$$

hence

$$\sum_j v_j \ln n_j = \sum_j v_j \ln c_j \rightsquigarrow \prod_j n_j^{v_j} = K(\tau) \quad (1.99)$$

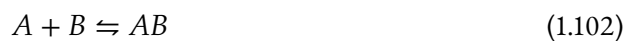
which is known as the **law of mass action**, defining

$$K(\tau) \equiv \prod_j c_j^{v_j} \exp\left(-\frac{v_j F_j^{\text{int}}}{\tau}\right) \quad (1.100)$$

Chemists often write for example

$$\frac{[A][B]}{[AB]} = K(\tau) \quad (1.101)$$

where $[\cdot]$ denotes the concentration. This would be for the reaction



I.7 PHASE TRANSITIONS

Under particular conditions two (or more) phases of a material can exist. However, if the system is at equilibrium the all the chemical potentials must be equal, for example

$$\mu_\ell(p, \tau) = \mu_g(p, \tau) \quad (1.103)$$

By expanding this linearly in p and τ we get that

$$\frac{dp}{d\tau} = \frac{\left(\frac{\partial \mu_\ell}{\partial \tau}\right)_p - \left(\frac{\partial \mu_g}{\partial \tau}\right)_p}{\left(\frac{\partial \mu_\ell}{\partial p}\right)_\tau - \left(\frac{\partial \mu_g}{\partial p}\right)_\tau} = \frac{s_g - s_\ell}{v_g - v_\ell} \quad (1.104)$$

note that $v = V/N$! The change in entropy is related to the heat transfer. We define the (additional) heat that is required for the phase transition as the **latent heat**:

$$L \equiv \tau(s_g - s_\ell) \quad (1.105)$$

Which means

$$\frac{dp}{d\tau} = \frac{L}{\tau \Delta v} \quad (1.106)$$

The **Clausius-Clapyron equation** (or **vapor pressure equation**).

Van der Waals equation of state

By modifying our expression for F for an ideal gas using the following substitution

$$V \rightsquigarrow V - Nb \quad (1.107)$$

where Nb accounts for the volume occupied by the particles and by accounting for their interactions using the *mean field method* we obtain

$$F_{\text{vdW}} = -N\tau \left(\ln \left(\frac{n_Q(V - Nb)}{N} \right) + 1 \right) - \frac{N^2 a}{V} \quad (1.108)$$

from which we can calculate the pressure $\left(\frac{\partial F}{\partial V}\right)_{\tau, N} = -p$, and by moving some terms around we obtain a modified ideal gas equation (**van der Waals equation of state**)

$$\left(p + \frac{N^2 a}{V^2} \right) (V - Nb) = N\tau \quad (1.109)$$

From here we define the following critical quantities

$$p_c = \frac{a}{27b^2}, \quad V_c = 3Nb, \quad \tau_c = \frac{8a}{27b} \quad (1.110)$$

Which allows us to write the **law of corresponding states**, which is a modified dimensionless gas equation

$$\hat{p} = \frac{\frac{8}{3}\hat{\tau}}{\hat{V} - \frac{1}{3}} - \frac{3}{\hat{V}^2} \quad (1.111)$$

where the $\hat{\cdot}$ implies we have divided the respective values by their corresponding critical values. Plotting $\hat{p}(\hat{V})$ yields

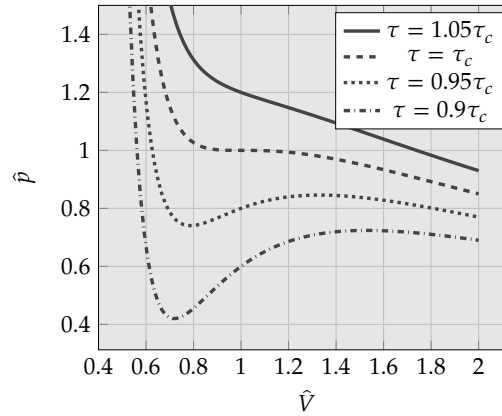


Figure 1.3: For $\tau < \tau_c$ the minimum of the function is not a point of inflection. For $\tau \geq \tau_c$ this cannot be said, which implies that there is no difference between the phases

If $\tau < \tau_c$ we can draw a *coexistence line* as in Figure in the book, along which the two phases can coexist. Once we've left the coexistence line there can only be one stable phase.

Nucleation: Our results can be used to predict the critical radius that is required for a droplet of liquid to be stable (droplets too small vaporise). Let

$$\Delta\mu = \mu_g - \mu_l \quad (1.112)$$

then

$$R_c = \frac{2\gamma}{n_l \Delta\mu} \quad (1.113)$$

where γ is the surface free energy and n_l is the concentration of the molecules in the liquid. For an example with ferromagnetism see pages 295-298

Landau theory of phase transitions

Suppose a system can be described by a single parameter, ξ , (as well as temperature), and hence we define the **Landau free energy function**

$$F_L(\xi, \tau) \equiv U(\xi, \tau) - \tau \sigma(\xi, \tau) \quad (1.114)$$

and let ξ_0 be the value of ξ at equilibrium. F_L is minimised at equilibrium, thus $F(\tau) = F_L(\xi_0, \tau) \leq F_L(\xi, \tau)$. Let us write out F_L in a power series

$$F_L(\xi, \tau) = g_0(\tau) + \frac{1}{2}g_2(\tau)\xi^2 + \frac{1}{4}g_4(\tau)\xi^4 + \dots \quad (1.115)$$

(we restrict ourselves to systems where F_L is even in ξ). The simplest example of a phase transition is a **second order phase transition** which is such if $g_2(\tau)$ changes sign at temperature τ_0 , for examples

$$F_L(\xi, \tau) = g_0(\tau) + \frac{1}{2}\alpha(\tau - \tau_0)\xi^2 + \frac{1}{4}g_4\xi^4 \quad (1.116)$$

We would like to minimise this in ξ , which is achieved at

$$\xi = 0 \vee \sqrt{\frac{\alpha(\tau_0 - \tau)}{g_4}} \quad (1.117)$$

(\vee is or). From this we get an expression for the Helmholtz free energy

$$F(\tau) = g_0(\tau) - \frac{\alpha^2}{4g_4}(\tau - \tau_0)^2 \quad (1.118)$$

First order transitions A first order transition is defined similarly to a second order transition, except that we additionally require that g_4 is negative:

$$F_L(\xi, \tau) = g_0(\tau) + \frac{1}{2}g_2(\tau)\xi^2 - \frac{1}{4}|g_4(\tau)|\xi^4 + \dots \quad (1.119)$$

First order phase transitions will have a very abrupt and usually discontinuous behaviour as a function of temperature, whereas second order transitions have a discontinuity in their first derivative—see pages 300-305.

I.A APPENDIX

Stirlings Approximation:

$$N! \approx (2\pi N)^{\frac{1}{2}} N^N \exp\left(-N + \frac{1}{12N} + \dots\right) \quad (1.120)$$

Which for large N yields

$$\ln N! \approx N \ln N - N + \frac{1}{2} \ln(2\pi N) \approx N \ln N - N \quad (1.121)$$

Approximating infinite sums as integrals:

The total energy of photons in a $L \times L \times L$ cavity is

$$U = \sum_n \langle \varepsilon_n \rangle = \sum_{n_x} \sum_{n_y} \sum_{n_z} \frac{\hbar \omega_n}{\exp(\hbar \omega_n) - 1}, \quad (1.122)$$

$$\omega_N = \frac{n\pi c}{L} \quad (1.123)$$

However, assuming each term in the sum is small we can approximate this triple sum by an integral

$$U \approx \frac{1}{8} \int_0^\infty \frac{4\pi n^2 \hbar \omega_n}{\exp(\hbar \omega_n/t) - 1} dn \quad (1.124)$$

The $1/8$ is because we are only looking at the positive octant (because n_x, n_y and n_z are positive) and the $4\pi n^2$ is because we are integrating the shell of a sphere, because the number of possibilities of n_x, n_y and n_z for which $n^2 = n_x^2 + n_y^2 + n_z^2$ goes like n^2 .

Fermion or Boson?

Particles with *integer* spin are bosons, whereas particles with *half-integer* spin are fermions, for example

e, p, n	Fermion
H_2	Boson
3He	Fermion
4He	Boson

For atoms: in general if the number of neutrons is odd the atom is a fermion, otherwise it's a boson.
(because each proton is matched by an electron).

MATHEMATICAL METHODS IN PHYSICS III

2.1 GROUPS

A group consists of a set and a composition rule, often denoted (G, \cdot) , where G is the set and \cdot is the composition rule. In order for (G, \cdot) to be a group it needs to fulfil the following axioms:

1. Associativity:

$$(g_a \cdot g_b) \cdot g_c = g_a \cdot (g_b \cdot g_c), \quad \forall g_a, g_b, g_c \in G \quad (2.1)$$

2. The identity, I , must exist, for which:

$$I \cdot g = g \cdot I, \quad I \in G, \quad \forall g \in G \quad (2.2)$$

3. Every element of the group must have an inverse

$$g \cdot g^{-1} = g^{-1} \cdot g = I, \quad \forall g \in G \quad (2.3)$$

It can be shown that every element has a *unique* inverse.

- (4.) Closure: If g is in the group and h is in the group then $g \cdot h$ is in the group.

Generally groups are *not* commutative, that is

$$g_a \cdot g_b \neq g_b \cdot g_a \quad (\text{generally}) \quad (2.4)$$

However, if this is the case we refer to the group as an **abelian group**. For example rotations in two dimensions form an abelian group under composition of rotations. However, rotations in three dimensions are not commutative – see Figure 2 in chapter I.1.

A subgroup, $H \subset G$, is group for which

$$h \in H \implies h \in G \quad (2.5)$$

Note that the axioms for groups put significant constraints on H . For example if G has N elements, and H has M elements, then $M|N$ (that means that N divided by M is an integer), this is **Lagrange's theorem**.

2.1.1 Cyclic subgroups

Given a finite group, G , let us pick an element, g , and repeatedly multiply it with itself. $g^2 \in G$ by closure, and hence $g^3 \in G$ etc. There must exist some positive integer, k , such that $G^k = I$,

because the group is finite. We have just created a cyclic subgroup of G , which has order k . So a cyclic subgroup of G is any subgroup that can be written in the form

$$\{I, g, g^2, \dots, g^{k-1}\}, \quad \text{for } 0 < k \leq |G| \quad (2.6)$$

where $|G|$ denotes the numbers of elements in G (the order of G).

2.1.2 Direct product of groups

Given two groups, (G, \cdot) and (H, \circ) , then the direct product $(G \otimes H, \bullet)$ is also a group. The elements of this group are

$$(g, h) \quad \text{for } g \in G, \quad h \in H \quad (2.7)$$

and multiplication of composition occurs elementwise:

$$(g_1, h_1) \bullet (g_2, h_2) = (g_1 \cdot g_2, h_1 \circ h_2) \quad (2.8)$$

The following is true:

$$|G \otimes H| = |G||H| \quad (2.9)$$

For example **Klein's Vierergruppe** is a direct product group $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. \mathbb{Z}_2 has elements $\{0, 1\}$ hence

$$\mathbb{Z}_2 \otimes \mathbb{Z}_2 = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \quad (2.10)$$

Under *elementwise addition*, modulo 2.

A. Zee is confusing and says that

$$\mathbb{Z}_2 \otimes \mathbb{Z}_2 = \left\{ \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \quad (2.11)$$

under *elementwise multiplication*. However, these two groups are in fact isomorphic to each other, and can therefore be thought of as the same group. In fact they are two representations of the same group structure.

2.1.3 Multiplication tables

Given a group of order n we can define its multiplication table, whose i - j -th entry is the composition $g_i \cdot g_j$. By the group axioms every element in the group appears *once and only once* in every column and every row. Additionally the first row and column are filled out trivially, because these are the elements that have been multiplied by the identity. Let's suppose our group has 4 elements, I, A, B and C :

	I	A	B	C
I	I	A	B	C
A	A	A^2	AB	AC
B	B	BA	B^2	BC
C	C	CA	CB	C^2

where the composition is implied. In the appendix I will out this table.

Abelian groups' multiplication tables are symmetric.

2.1.4 Homomorphism and Isomorphism

A map $\Phi : G \rightarrow G'$ is a **homomorphism** if

$$\Phi(g_1 g_2) = \Phi(g_1) \Phi(g_2), \quad g_1, g_2 \in G \quad (2.12)$$

This mapping is an *isomorphism* if it additionally is bijective. For example $\mathbb{Z}_6 \cong \mathbb{Z}_2 \otimes \mathbb{Z}_4$, where " \cong " means "is isomorphic to". But A. Zee doesn't write \cong , instead he uses $=$, which I find very confusing.

2.1.5 Important fact about groups

Every finite group is isomorphic to a subgroup of S_n for some n . Think about the multiplication table, and that "title"-row is the unpermuted group (it's just the group in some order). The first row (where we multiplied by I) is thus just the trivial permutation, the second row is some permutation etc. Therefore each element in the the group corresponds to some permutation of the group – thus we can define an isomorphism from every group to a subgroup of S_n .

2.1.6 Finite Groups

The **permutation group**, S_n , describes the permutation of n numbers (or letters or objects). There are different ways of denoting the elements in S_n . Let us look at S_5 specifically. An element in S_5 might be

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix} \quad (2.13)$$

where this means that $1 \rightsquigarrow 4, 2 \rightsquigarrow 1, 3 \rightsquigarrow 5, 4 \rightsquigarrow 2$ and $5 \rightsquigarrow 3$. One can also use cycle notation. Our example from before would be written

$$g = (142)(35) \quad (2.14)$$

because 1 goes to 4 which goes to 2 which goes to 1 and 3 goes to 5 which goes to 3.

Any permutation in S_n can be written as a product of n_j j -cycles, provided $\sum_j j n_j = n$. Given a cycle structure characterised by integers n_j , then there are

$$\mathcal{N} = \frac{n!}{\prod_j j^{n_j} n_j!} \quad (2.15)$$

cycles with this structure.

A_n is the subgroup of S_n that contains all of S_n s *even* permutation. Dimensions of these

	d
S_n	$n!$
A_n	$\frac{n!}{2}$

2.1.7 Rules for multiplying cycles:

It's always easiest just to write it out. For example what is $(135)(24)(521)(34)$? Well

$$(135)(24)(521)(34) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \end{pmatrix} \quad (2.16)$$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix} = (1)(23)(45) = (23)(45) \quad (2.17)$$

However there are tricks that can be used so one doesn't need to write it all out.

1. If two cycles do not have any numbers in common, then they commute.
2. If two cycles have one number in common then they can be contracted, for example $(12)(23) = (123)$. Also $(12)(23)(34) = (12)(234) = (1234)$
3. $(12)(21) = I$. Because swapping two things twice is the same as not doing anything

A two-cycle (cycle with two numbers in it) is *odd* and a three-cycle is even. Generally even numbered cycles are odd and odd numbered cycles are even. Another way of saying this is that a permutation is even if it decomposes into an even number of 2-cycles (and odd otherwise).

2.1.8 Equivalence Classes

Two elements, g and g' , in a group G are equivalent (and hence in the same equivalence class) if

$$g' = f^{-1}gf \quad (2.18)$$

for some $f \in G$. Equivalence is transitive, thus if g and g' are equivalent and g' and g'' are equivalent then g and g'' are too.

1. In an abelian group every element is only equivalent to itself
2. The identity is always in its own equivalence class
3. The inverse of all elements in a class also form a class

2.1.9 Dihedral group D_n

The dihedral group describes rotations together with reflections of an n -sided polygon. Let R be the rotation through $2\pi/n$ radians and r the reflection through a *median* (if n is odd the median

passes through the centre of the polygon and one vertex, if n is even the median passes through two vertices, each on opposite sides). The group can be represented by

$$D_n = \{I, R, R^2, \dots, R^{n-1}, r, Rr, R^2r, \dots, R^{n-1}r\} \quad (2.19)$$

note that

$$R^n = I, \quad r^2 = I \quad (2.20)$$

2.1.10 The quaternionic group Q

Hamilton generalised the imaginary unit with his quaternions, they are defined as

$$i^2 = j^2 = k^2 = -1 \quad (2.21)$$

$$ij = -ji = k \quad (2.22)$$

$$jk = -kj = i \quad (2.23)$$

$$ki = -ik = j \quad (2.24)$$

which I suppose could be written as

$$\alpha\beta = \varepsilon_{\alpha\beta\gamma}\gamma - \delta_{\alpha\beta} \quad (2.25)$$

This group is not abelian (it's anticommutative).

2.1.11 The invariant subgroup

Let H be a subgroup of G . H is an **invariant subgroup** if

$$ghg^{-1} \in H, \quad \forall h \in H, \forall g \in G \quad (2.26)$$

This is sometimes also written as

$$gHg^{-1} = H, \quad g \in G \quad (2.27)$$

hence the subgroup H is invariant under similarity transformations.

2.1.12 Derived subgroups

Given a group, G , then the following is the **derived subgroup**:

$$\{a^{-1}b^{-1}ab : a, b \in G\} \quad (2.28)$$

So we look at all the products $a^{-1}b^{-1}ab$ and let a and b range through all elements in G .

2.1.13 Rotations

We can expand infinitesimal rotations up to the linear term, that is

$$R(\theta) \approx I + A \quad (2.29)$$

where A is an infinitesimal matrix of order θ . Rotations must be orthogonal

$$R^T R = I = (I + A^T)(I + A) = I + A^T + A = I \quad (2.30)$$

again neglecting higher order terms, hence $A^T = -A$. If A is antisymmetric, then it is proportional to

$$\mathcal{J} \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (2.31)$$

hence

$$R(\theta) \approx I + \theta \mathcal{J} = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix} \quad (2.32)$$

The full matrix is given by infinitely many infinitesimal rotations:

$$R(\theta) = \lim_{N \rightarrow \infty} \left(R\left(\frac{\theta}{N}\right) \right)^N = e^{\theta \mathcal{J}} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (2.33)$$

This approach can also be generalised to higher dimensions, however our matrix \mathcal{J} is replaced by numerous matrices (in d dimensions there are $\frac{1}{2}d(d-1)$ of these). For example in three dimensions

$$\mathcal{J}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathcal{J}_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.34)$$

In which case

$$R(\theta) = e^{\theta_x \mathcal{J}_x + \theta_y \mathcal{J}_y + \theta_z \mathcal{J}_z} \quad (2.35)$$

But physicists often define $J_k = -i \mathcal{J}_k$, hence

$$R(\theta) = e^{i \sum_k \theta_k J_k} \quad (2.36)$$

The matrices J are called the **generators** of the rotation group (and the Lie algebra). For higher dimensions we must use two indices to indicate which rotation is meant. For example J_{13} would be rotating the 1-3-plane about itself. The ab -th entry in the matrix J_{mn} is given by

$$(J_{mn})^{ab} = -i (\delta^{ma} \delta^{nb} - \delta^{mb} \delta^{na}) = -i \epsilon^{abc} \epsilon^{mnc} \quad (2.37)$$

this holds in all dimensions. These are the generators of $SO(N)$, the *special orthogonal group*. We denote it's Lie algebra as $so(N)$. "Special" refers to the fact that the matrices have $\det = 1$ and "orthogonal" refers to the orthogonality of all the group's elements. N specifies that the groups describes rotations in N dimensions.

2.1.14 Lie Algebra

Generally rotations do not commute (again see Figure 2 in chapter I.1.), hence we define the *commutator*:

$$[A, B] \equiv AB - BA \quad (2.38)$$

In the case of the generators

$$[J_a, J_b] = \sum_c i\epsilon_{abc} J_c \quad (2.39)$$

where ϵ_{abc} is the totally antisymmetric tensor (Levi-Civita). ϵ_{abc} is the **structure constant** of the Lie Algebra. Generally, however,

$$[T_a, T_b] = \sum_c if_{abc} T_c = if_{abc} T_c \quad (2.40)$$

where T_i are the generators of the Lie Algebra. Due to the fact that $[\cdot, \cdot]$ is antisymmetric (by definition), f_{abc} is also antisymmetric. Note that I have used Einsteins summation convention, I will try to avoid this at all costs to make things easier to read.

2.2 REPRESENTATION THEORY & CHARACTER TABLES

The notion of a group and an algebra is quite abstract – the elements in these *can* be vectors and matrices, but the can be anything. However, representation theory is all about representing these elements my matrices or tensors. Let $D(g)$ be the matrix that represents the group element g , then it must be true that

$$D(g_1 g_2) = D(g_1) D(g_2) \quad (2.41)$$

Note that this is the definition of a homomorphism! Usually we deal with isomorphisms, because it makes sense to allocate a unique matrix to every group element.

At this stage there is a bit of confusion. The group elements are represented by matrices (and later on tensors), and these matrices act on vectors (lower rank tensors later on). For example rotations in 2D can be represented by

$$R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (2.42)$$

and these matrices act upon (and rotate) 2-vectors. We say that this representation is *furnished by vectors*.

Now that we have defined the representation, we can define the **character**, χ , of each element, g :

$$\chi^{(r)}(g) = \text{tr } D^{(r)}(g), \quad g \in G \quad (2.43)$$

The character is dependent on which representation we chose, which is why we need to include the superscript (r) , which just tells us which representation we are looking at.

An important characteristic of the trace is that it is invariant under cyclic permutation, hence

$$\text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA) \quad (2.44)$$

Therefore every element in an equivalence class has the character:

$$\text{tr}(f^{-1}gf) = \text{tr}(ff^{-1}g) = \text{tr}(g) \quad (2.45)$$

Similarity transformations of any representation does not change the character:

$$D' = S^{-1}DS \rightsquigarrow \text{tr } D' = \text{tr } S^{-1}DS = \text{tr } D \quad (2.46)$$

2.2.1 Reducible and irreducible representations

Let's look at $SO(3)$. The trivial 1-dimensional representation ($D^{(1)}(g)$) is a representation of all groups (it fulfils all the requirements). Another representation of $SO(3)$ is the 3-dimensional representation ($D^{(3)}(g)$), which we know well (3×3 special orthogonal matrices). We can take the direct sum of these representations, for example

$$D(g) = D^{(1)}(g) \oplus D^{(1)}(g) \oplus D^{(3)}(g) \oplus D^{(3)}(g) \quad (2.47)$$

$$= \left(\begin{array}{c|c|c|c} D^{(1)}(g) & 0 & 0 & 0 \\ \hline 0 & D^{(1)}(g) & 0 & 0 \\ \hline 0 & 0 & D^{(3)}(g) & 0 \\ \hline 0 & 0 & 0 & D^{(3)}(g) \end{array} \right) \quad (2.48)$$

this is a 8-dimensional representation, yet it is not particularly interesting, because it decomposes into two 1-dimensional representations and two 3-dimensional representations. Therefore we say that this representation is **reducible**. We focus primarily on **irreducible** representations, as we can build all reducible representations using these.

Unitarity Theorem: All finite groups have unitary representations, that is representations for which

$$D(g)^\dagger D(g) = I, \quad \forall g \in G \quad (2.49)$$

In general this is *not* true for infinite groups. However if it does hold for a infinite group, we refer to the group as **compact**.

Schur's Lemma: If $D(g)$ is an irreducible representation of a finite group, G , and if there is some matrix, A , such that $AD(g) = D(g)A$ for all g , then $A = \lambda I$ for some constant λ .

That is, if $D(g)$ is a irreducible representation of a group, then *the only matrix that commutes with* $D(g)$ is (a constant times) the identity matrix.

2.2.2 The Great Orthogonality Theorem:

Given a d -dimensional irreducible representation $D(g)$ of a finite group G , we have

$$\sum_g D^\dagger(g)_j^i D(g)_l^k = \frac{N(G)}{d} \delta_l^i \delta_j^k \quad (2.50)$$

Additionally for two representations r and s we have that

$$\sum_g D^{(r)\dagger}(g)_j^i D^{(s)}(g)_l^k = \frac{N(G)}{d_r} \delta_r^s \delta_l^i \delta_j^k \quad (2.51)$$

Let $i = j$ and $k = l$ and sum over all values of i and k , then we get:

$$\sum_{g,i,k} D^{(r)\dagger}(g)_i^i D^{(s)}(g)_k^k = \sum_{i,k} \frac{N(G)}{d_r} \delta_r^s \delta^{ik} \delta^{ki} \quad (2.52)$$

The left hand side is the product of the traces of the two matrices and hence the product of the characters in each representation, and the right hand side gives us d_r times what we had before (because i and k are summed over d_r different values), thus

$$\sum_c n_c (\chi^{(r)}(c))^* \chi^{(s)}(c) = N(G) \delta^{rs} \quad (2.53)$$

the n_c appears because we change our summation from going over each group element, to going over each class. The class, c , has n_c elements.

Now we are ready to introduce the **character table**, begin by writing out each equivalence class, each in its own row of the table. Next write add a column for each irreducible representation. Note that

$$N(R) = N(C) \quad (2.54)$$

The number of irreducible representations is equal to the number of equivalence classes. Hence the table has as many rows as it has columns – it is square. Now we fill in the elements: the element in row c and column r is equal to $\chi^{(r)}(c)$. For example, for A_4 :

A_4	n_c	c	1	1'	1''	3
	1	I	1	1	1	3
	3	$(12)(34)$	1	1	1	-1
	4	(123)	1	ω	ω^*	0
	4	(132)	1	ω^*	ω	0

We can find ω by using the different rules we just discussed. The result is

$$\omega = e^{i\frac{2\pi}{3}} \quad (2.55)$$

In the appendix I fill out the character table for Q .

Using these we can also **test for reducibility**: Given a representation $D(g)$ its characters will be given by

$$\chi(c) = \sum_r n_r \chi^{(r)}(c) \quad (2.56)$$

For r ranging over all irreducible representations (This is naturally true – if $D(g)$ is irreducible then n_r is zero for all representations, except when it is one. If $D(g)$ is reducible it must decompose into irreducible representations, in which case this is most definitely true). It can be shown that

$$\sum_c n_c \chi^*(c) \chi(c) = N(G) \sum_r (n_r)^2 \quad (2.57)$$

hence by taking the trace of our representation (and summing over all classes) we can establish which irreducible representations appear how many times in the direct sum!

2.2.3 Regular representation

Let $N(G)$ be the number of elements of the group (i.e the order of the group). As discussed previously, every group is isomorphic to a subgroup of $S_{N(G)}$, and $S_{N(G)}$ has a defining representation

with dimension $N(G)$. Hence every group has a representation of order $N(G)$, which is known as the **regular representation**. Note that the regular representation is reducible (usually). For the regular representation $n_r = d_r$, hence the regular representation decomposes into n_r terms of the representation r . The regular representation is block diagonal.

Here are all the rules we have for the character tables:

1. Dimensions of the irreducible representations

$$\sum_r d_r^2 = N(G) \quad (2.58)$$

2. Column orthogonality

$$\sum_c n_c (\chi^{(r)}(c))^* \chi^{(s)}(c) = N(G) \delta^{rs} \quad (2.59)$$

3. Row orthogonality

$$\sum_r (\chi^{(r)}(c))^* \chi^{(r)}(c') = \frac{N(G)}{n_c} \delta^{cc'} \quad (2.60)$$

4. The character table is square

$$N(C) = N(R) \quad (2.61)$$

Additionally we have two rules for the decomposition of reducible representations

1. $\sum_c n_c \chi^*(c) \chi(c) = N(G) \sum_r n_r^2$
2. $\sum_c n_c (\chi^{(r)}(c))^* \chi(c) = N(G) n_r$

2.2.4 Reality

Representations fall into three groups: **real**, **pseudoreal** and **complex**. Real and pseudoreal representations have (purely) real characters, whereas complex representations have complex characters. This discussion begins when we realise that if $D(g)$ is a representation, then so is $D^*(g)$. The characters of $D^*(g)$:

$$\chi^{(r^*)}(g) = \text{tr } D(g)^* = \text{tr } D(g)^* = (\chi^{(r)}(g))^* \quad (2.62)$$

We can check whether a given representation is real, using the **reality checker**:

$$\sum_{g \in G} \chi^{(r)}(g^2) = \eta^{(r)} N(G), \quad \text{with } \eta^{(r)} = \begin{cases} 1 & \text{if real} \\ 0 & \text{if complex} \\ -1 & \text{if pseudoreal} \end{cases} \quad (2.63)$$

Note that if a representation is for example real, it does not necessarily mean that all entries in the matrix are real, but rather that its trace is real (and hence there exists a similarity transformation that "makes" a matrix that only has real entries).

For the case of pseudoreality it is true that

$$D(g)^* = SD(g)S^{-1} \quad (2.64)$$

i.e there exists a similarity transformation that takes you from $D(g)$ to $D(g)^*$.

2.2.5 Number of square roots

Starting with the reality checker note that if an element f has σ_f square roots then it will appear σ_f times in the sum, hence

$$\sum_f \sigma_f \chi^{(r)}(f) = \eta^{(r)} N(G) \quad (2.65)$$

From here we can show that

$$\sigma_f = \sum_r \eta^{(r)} \chi^{(r)}(f) \quad (2.66)$$

Remember the sum on the right only goes over the irreducible representations (of course, otherwise there would be infinite terms)

2.3 ROTATION GROUPS AND TENSORS $\mathbf{SO}(3)$

For $SO(3)$ we don't need to worry about the difference between upper and lower indices, therefore in this entire section we will only use upper indices.

The representation of $SO(3)$ that we are best acquainted with is the regular representation, i.e 3×3 (special orthogonal) matrices. These matrices operate on (and rotate) vectors, that is the vector element V^i is transformed as follows:

$$V^i \rightsquigarrow V'^i = \sum_j R^{ij} V^j \quad (2.67)$$

where i and j take on three different values. We say that our regular representation is *furnished* by vectors. Similarly for $SO(N)$ the rotation matrix is $N \times N$ and the indices take on N different values. However, this is not the most general form. Here we introduce the **tensor**:

$$T^{ij} \rightsquigarrow T'^{ij} = \sum_{kl} R^{ik} R^{jl} T^{kl} \quad (2.68)$$

The representation that we get from studying this transformation is furnished by rank-2 tensors. Note that in the case of $SO(3)$ the indices can take on three values, hence the representation is 9-dimensional.

Rank-2 means that the tensor has 2 indices.

But are these representations reducible? It turns out there are two types of tensors that are transform independently from one another under these transformations:

1. Traceless symmetric tensors
2. Antisymmetric tensors

This can be seen by defining

$$A^{ij} = T^{ij} - T^{ji} \quad (2.69)$$

this transforms like

$$A^{ij} \rightsquigarrow A'^{ij} = T'^{ij} - T'^{ji} \quad (2.70)$$

$$= \sum_{kl} R^{ik} R^{jl} T^{kl} - \sum_{kl} R^{ik} R^{jl} T^{lk} \quad (2.71)$$

$$= \sum_{kl} R^{ik} R^{jl} (T^{kl} - T^{lk}) \quad (2.72)$$

$$= \sum_{kl} R^{ik} R^{jl} A^{kl} \quad (2.73)$$

A^{ij} as an antisymmetric tensor. Next consider

$$S^{ij} = T^{ij} + T^{ji} \quad (2.74)$$

This transforms like

$$S^{ij} \rightsquigarrow S'^{ij} = T'^{ij} + T'^{ji} \quad (2.75)$$

$$= \sum_{kl} R^{ik} R^{jl} T^{kl} + \sum_{kl} R^{ik} R^{jl} T^{lk} \quad (2.76)$$

$$= \sum_{kl} R^{ik} R^{jl} (T^{kl} + T^{lk}) \quad (2.77)$$

$$= \sum_{kl} R^{ik} R^{jl} S^{kl} \quad (2.78)$$

Note that if $i = j$ we get

$$\sum_i S^{ii} \rightsquigarrow \sum_i S'^{ii} = \sum_{kl} R^{ik} R^{il} S^{kl} \quad (2.79)$$

$$= \sum_{kl} (R^T)^{ki} R^{il} S^{kl} \quad (2.80)$$

$$= \sum_{kl} (R^{-1})^{ki} R^{il} S^{kl} \quad (2.81)$$

$$= \sum_{kl} \delta^{kl} S^{kl} = \sum_k S^{kk} \quad (2.82)$$

Hence the trace is invariant under these transformations. Thus a symmetric tensor transforms as two independent parts, a traceless symmetric tensor and a trace. Therefore we tend to work with symmetric traceless tensors.

Dimensions of rank-2 tensors: The first is easy: the trace of a rank-2 tensor is a number, hence 1-dimensional. The antisymmetric tensors have $N - 1$ numbers in the first row, $N - 2$ in the second all the way down until there are 0 numbers left.

$$\sum_{n=0}^{N-1} n = \frac{1}{2}N(N - 1) \quad (2.83)$$

thus antisymmetric rank-2 tensors are $\frac{1}{2}N(N - 1)$ -dimensional. Symmetric rank-2 tensors have the same number *plus* the diagonal; there are N numbers on the diagonal

$$\frac{1}{2}N(N - 1) + N = \frac{1}{2}N(N + 1) \quad (2.84)$$

thus symmetric rank-2 tensors are $\frac{1}{2}N(N+1)$ -dimensional

Therefore our 9×9 matrix from before (the regular representation of $SO(3)$) decomposes into an antisymmetric part, a symmetric traceless part and a trace, hence there exists a similarity transformation such that

$$S^{-1}D(R)S = \left(\begin{array}{c|c|c} 3 \times 3 \text{ block} & 0 & 0 \\ \hline 0 & 1 \times 1 \text{ block} & 0 \\ \hline 0 & 0 & 5 \times 5 \text{ block} \end{array} \right) \quad (2.85)$$

2.3.1 Rank- n tensors in $SO(N)$

To generalise what we did before for rank-2 tensors, we will look at

$$T^{i_1 i_2 \dots i_n} \quad (2.86)$$

where each of the indices i_1, i_2, \dots, i_n can take on values between 1 and N . These tensors transform very similarly to rank-2 tensors:

$$T^{i_1 i_2 \dots i_n} \rightsquigarrow T'^{i_1 i_2 \dots i_n} \quad (2.87)$$

$$= \sum_{j_1, j_2, \dots, j_n} R^{i_1 j_1} R^{i_2 j_2} \dots R^{i_n j_n} T^{j_1 j_2 \dots j_n} \quad (2.88)$$

As before these split into symmetric traceless tensors, antisymmetric tensor as well as a trace. However, note that now the trace is a rank- $(n-2)$ tensor. This brings us to the **contraction of indices**, and **dual tensors**. Contraction of indices occurs by using the Kronecker delta:

$$\sum_{i_\alpha, i_\beta} \delta^{i_\alpha i_\beta} T^{i_1 \dots i_n} = \sum_{i_\alpha, i_\beta} T^{i_1 \dots i_\alpha \dots i_\alpha \dots i_n} \rightsquigarrow \quad (2.89)$$

$$\sum_{i_\alpha, i_\beta} T'^{i_1 \dots i_\alpha \dots i_\alpha \dots i_n} \quad (2.90)$$

$$= \sum_{j_\alpha, j_\beta} R^{i_1 j_1} \dots R^{i_\alpha j_\alpha} \dots R^{i_\alpha j_\beta} \dots R^{i_n j_n} T^{j_1 \dots j_\alpha \dots j_\alpha \dots j_n} \quad (2.91)$$

$$= \sum_{j_\alpha \text{ does not include } R^{i_\alpha j_\alpha} \text{ nor } R^{i_\beta j_\beta}} \underbrace{R^{i_1 j_1} \dots R^{i_n j_n}}_{\text{unchanged}} T^{j_1 \dots j_\alpha \dots j_\alpha \dots j_n} \quad (2.92)$$

In the final equality I used, as before, that: $R^{i_\alpha j_\alpha} R^{i_\alpha j_\beta} = \delta^{j_\alpha j_\beta}$. Hence by renaming the indices k_1, k_2, \dots, k_{n-2} it becomes clear that the Kronecker delta has contracted two indices (i_α and i_β) such that we are now left with a rank- $(n-2)$ tensor.

We create the **dual tensor** by using the totally antisymmetric tensor on an antisymmetric tensor, $A^{i_1 i_2}$ (this is not the general case, but it's useful to start with this). Note that the totally antisymmetric tensor in $SO(N)$ will have N indices. Remember we have to sum over repeated indices, as usual thus

$$\sum_{i_1, i_2} \varepsilon^{i_1 i_2 \dots i_N} A^{i_1 i_2} = B^{j_1 j_2 \dots j_{N-2}} \quad (2.93)$$

thus $A^{i_1 i_2}$'s dual tensor is $B^{j_1 j_2 \dots j_{N-2}}$. Generally if we begin with a rank- p tensor, using ε on it will result in a rank- $(N-p)$ tensor.

2.3.2 Dimensions of rank- n tensors, whose indices can take on N values

As usual we must differentiate between antisymmetric tensors, symmetric traceless tensors and traces.

2.3.3 Antisymmetric

An antisymmetric tensor with n indices that can take on N values has the following dimension:

$$d = \frac{N!}{n!(N-n)!} = \binom{N}{n} \quad (2.94)$$

2.3.4 Symmetric Traceless

A symmetric tensor with n indices that can take on N values has the following dimension:

$$d = \frac{(N+n-1)!}{n!(N-1)!} - \frac{(N+n-3)!}{(n-2)!(N-1)!} \quad (2.95)$$

$$= \binom{N+n-1}{n} - \binom{N+n-3}{n-2} \quad (2.96)$$

2.3.5 $SO(3)$

We often look at $SO(3)$ specifically. First note that for $SO(3)$ it's only important to look at symmetric tensors (we can look at the dual of antisymmetric tensors, which generally has fewer indices, for example an antisymmetric tensor with 2-indices can be written as a tensor with one index). Remember A. Zee's inductive approach to solving this problem. Thus let us look at a symmetric traceless tensor in $SO(3)$ with j indices:

$$d = \frac{(j+2)!}{j! 2!} - \frac{j!}{(j-2)! 2!} \quad (2.97)$$

$$= \frac{1}{2} ((j+1)(j+2) - j(j-1)) = 2j+1 \quad (2.98)$$

*Spin, angular momentum
and probably other
quantum numbers...*

which is related to there being $2j+1$ states in a spin- j system!

2.3.6 Adjoint representation

The **adjoint representation** is defined as

$$(T^a)^{bc} = -if^{abc} \quad (2.99)$$

That is, the matrix T^a 's b - c -th element is given by a constant times the structure constant. For example, the matrices J from Equation 2.36 are the adjoint representation of $SO(3)$.

It can be shown that the adjoint representation indeed is a representation by using Jacobi's relation, which states that

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \quad (2.100)$$

which, using the definition of our Lie Algebra can be written as

$$f^{abd}f^{dca} + f^{bcd}f^{dab} + f^{cad}f^{adb} = 0 \quad (2.101)$$

Now plugging our definition of the adjoint representation we can show that this holds.

2.3.7 Ladder operators and $\mathfrak{so}(3)$

We would like to represent $\mathfrak{so}(3)$ with matrices; what this means is that we find matrices J_i such that

$$[J_a, J_b] = \sum_c i\epsilon_{abc}J_c \quad (2.102)$$

We know from before that there exists a $(2j + 1)$ -dimensional representation of $SO(3)$, for j non-negative. Hence there must exist a $(2j + 1)$ -dimensional representation of the Lie Algebra. We introduce the ladder operators (raising and lowering operators), which will help us determine the entries in matrices J_i :

$$J_{\pm} \equiv J_x \pm iJ_y \quad (2.103)$$

note that

$$[J_z, J_{\pm}] = \pm J_{\pm}, \quad [J_+, J_-] = 2J_z \quad (2.104)$$

By using smart equations involving these commutation relations we can show that

$$J_{\pm}|m\rangle = c_{m\pm 1}|m \pm 1\rangle = \sqrt{(j + 1 \pm m)(j \mp m)}|m \pm 1\rangle \quad (2.105)$$

Using these coefficients and inverting the equations in Equation 2.103 we can obtain an expression for J_x and J_y . Note that $|m\rangle$ is in J_z 's eigenbasis, hence

$$J_z|m\rangle = m|J_z\rangle \quad (2.106)$$

Hence it is diagonal with the values of m on the diagonal. Remember the equation from Quantum Mechanics

$$Q_{ij} = \langle i|\hat{Q}|j\rangle \quad (2.107)$$

These coefficients terminate for $m = \pm j$. That is $J_+|j\rangle = 0$ and $J_-|-j\rangle = 0$, hence there are $2j + 1$ states, as expected.

2.3.8 Multiplying two representations together (still $SO(3)$)

In general multiplying two irreducible representations together results in a reducible representation. Let us begin with totally symmetric traceless tensors, $S^{i_1 i_2 \dots i_j}$ and $T^{k_1 k_2 \dots k_j}$. Now we begin by creating a new tensor

$$P^{i_1 \dots i_j k_1 \dots k_j} = S^{i_1 i_2 \dots i_j} T^{k_1 k_2 \dots k_j} \quad (2.108)$$

which only has *separate* symmetry properties for its first j indices and its final j' indices. We begin by symmetrising this and removing the trace, which gives us a rank- $(j + j')$ tensor. Now we use ε^{ikl} on P . Note that i and k are respectively an index from S and T (they can't *both* be from either S or T , because these are symmetric by construction, and using ε on a symmetric tensor gives you zero). Thus we trade off two indices for one, giving us a new tensor with $(j + j' - 1)$ indices. We symmetrise and remove the trace, giving us a traceless symmetric rank- $(j + j' - 1)$ tensor. This is repeated until one of the tensors S or T runs out of indices, thus

$$j \otimes j' = (j + j') \oplus (j + j' - 1) \oplus \cdots \oplus |j - j'| \quad (2.109)$$

$|j - j'|$ gives us the number of indices that remains on the tensor that has most indices. We know the dimension of each term on the right side (for example $2(j + j') + 1$), and by adding these dimensions we can determine the dimension of the coupled system (direct product system):

$$d = \sum_{k=|j-j'|}^{j+j'} k = (2j + 1)(2j' + 1) \quad (2.110)$$

Let us look at two spin- $\frac{1}{2}$ systems, as we did in Quantum Mechanics. For these $j = j' = \frac{1}{2}$, hence

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0, \quad (j \text{ and } j') \quad (2.111)$$

I have added a parenthesis that says j and j' because the numbers in the equation are the values of j and j' (and linear combinations). In a second we will write a similar equation where the numbers in the instead represent the dimension of the representation

which means that when you couple two spin- $\frac{1}{2}$ systems you get a spin-1 and a spin-0 system! Recall from QM when we looked at the spin-spin coupling of an electron and proton of hydrogen in its ground state we saw that the four states $\left(\frac{1}{2} \cdot 2 + 1\right)^2 = 4$ split into a triplet (a spin-1 system has 3 states) and a singlet (spin-0 system has one state).

2.3.9 How do we name irreducible representation?

Instead of denoting our representation by j , we can represent it by its dimension. Hence Equation 2.111 becomes

$$2 \otimes 2 = 3 \oplus 1, \quad (\text{dimensions}) \quad (2.112)$$

Which is nice because $2 \cdot 2 = 3 + 1$. This works only when are looking at the dimensions of the representations, not the values of j and j' .

2.3.10 But which states?– Clebsch-Gordon decomposition

So we have seen that $j \otimes j'$ decomposes into a direct sum of systems with j'' between $|j - j'|$ and $j + j'$, and hence each one of these systems has between $2|j - j'| + 1$ and $2(j + j') + 1$ states. But which states go where and what are the states actually?

Well, the systems with spin- j and spin- j' have the states

$$|j, m\rangle, \quad |j', m'\rangle \quad (2.113)$$

respectively. We are taking the tensor product of these two systems, so why not denote the coupled system's states with

$$|j, m\rangle \otimes |j', m'\rangle \quad (2.114)$$

Griffiths writes $|j, j', m, m'\rangle$, but it's the same thing – what we called the *coupled basis* in QM2. We are about to find the uncoupled basis. It can be shown that the product rule holds for the operators operating on tensor products of states, which means that, for example

$$J_z(|j, m\rangle \otimes |j', m'\rangle) \quad (2.115)$$

$$= (J_z|j, m\rangle) \otimes |j', m'\rangle + |j, m\rangle \otimes (J_z|j', m'\rangle) \quad (2.116)$$

$$= (m + m')(|j, m\rangle \otimes |j', m'\rangle) \quad (2.117)$$

So let us do something smart. We start with the state $|j, j\rangle \otimes |j', j'\rangle$, and use the lowering operator:

$$J_- (|j, j\rangle \otimes |j', j'\rangle) \quad (2.118)$$

$$= c_{j-1}|j, j-1\rangle \otimes |j', j'\rangle + c_{j'-1}|j, j\rangle \otimes |j', j'-1\rangle \quad (2.119)$$

And now we take that state and use the lowering operator once again. This will give us $2(j + j') + 1$ states – the states in the representation we called $(j + j')$. But this is not all the states in the system! we are expecting $(2j + 1)(2j' + 1)$. So let us find a state that is orthogonal to $J_- (|j, j\rangle \otimes |j', j'\rangle)$ and repeat the procedure again. We do this until we have all the states. For example, if $j = j' = \frac{1}{2}$ we get

$$|\frac{1}{2}, \frac{1}{2}\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle \quad (2.120)$$

$$\frac{1}{\sqrt{2}} \left(|\frac{1}{2}, -\frac{1}{2}\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle + |\frac{1}{2}, \frac{1}{2}\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle \right) \quad (2.121)$$

$$|\frac{1}{2}, -\frac{1}{2}\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle \quad (2.122)$$

$$\frac{1}{\sqrt{2}} \left(|\frac{1}{2}, -\frac{1}{2}\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle - |\frac{1}{2}, \frac{1}{2}\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle \right) \quad (2.123)$$

Four states, just as expected. The first three states are the states in the spin-1 system and the last state is the spin-0 system's state.

2.4 SPECIAL UNITARY GROUPS $SU(N)$

The special unitary groups $SU(N)$ consist of $N \times N$ matrices, U , that are unitary, i.e for which it is true that

$$U^\dagger U = I, \quad \det(U) = 1 \quad (2.124)$$

A few important facts about $SU(N)$:

1. Now it is important whether we write indices on top of at the bottom. For example $\phi^i = (\phi_i)^*$
2. When we use the repeated index summation rule this only is true if one of the indices is on top and the other is at the bottom, that is

$$\sum_i \phi_i \psi^i \quad (2.125)$$

3. The kronecker delta now has a lower and an upper index:

$$\delta_i^j \quad (2.126)$$

4. The totally antisymmetric tensor comes in to forms

$$\varepsilon^{i_1 i_2 \dots i_n}, \quad \varepsilon_{i_1 i_2 \dots i_n} \quad (2.127)$$

Now, vector transformations look like

$$\psi^i \rightsquigarrow \psi'^i = \sum_j U_j^i \psi^j \quad (2.128)$$

Similarly we can define transformations of covectors

$$\psi_i \rightsquigarrow \psi'_i = \sum_j (U^\dagger)_i^j \psi_j \quad (2.129)$$

Thus tensors can have upper and lower indices, and that these indices should be thought of as separate entities. Generally tensors transform like

$$T_{j_1 j_2 \dots j_n}^{i_1 i_2 \dots i_n} \rightsquigarrow T_{j'_1 j'_2 \dots j'_n}^{i'_1 i'_2 \dots i'_n} \quad (2.130)$$

$$= \sum_{l_1 l_2 \dots l_n} \sum_{k_1 k_2 \dots k_n} U_{k_1}^{i'_1} \dots U_{k_n}^{i'_n} T_{l_1 l_2 \dots l_n}^{k_1 k_2 \dots k_n} (U^\dagger)_{j'_1}^{l_1} \dots (U^\dagger)_{j'_n}^{l_n} \quad (2.131)$$

But nobody can read that, so let's take an example: The tensor T_{klm}^{ij} transforms like

$$T_{klm}^{ij} \rightsquigarrow \quad (2.132)$$

$$T_{klm}^{'ij} = \sum_{nopqr} U_n^i U_o^j T_{pqr}^{no} (U^\dagger)_k^p (U^\dagger)_l^q (U^\dagger)_m^r \quad (2.133)$$

The rules from before about contraction of indices and dual tensors are *similar* to how they were in $SO(3)$ however, slightly more complicated. Now we can:

1. Contract an upper with a lower index using δ_k^i . So that means out of a tensor with j upper indices and j' lower indices we make a tensor with $(j - 1)$ upper indices and $(j' - 1)$ lower indices. I'll write it out for a specific case:

$$\sum_{i,k} \delta_k^i F_k^{ij} = \sum_i F_i^{ij} = G^i \quad (2.134)$$

so $\sum_{i,k} \delta_k^i F_k^{ij}$ transforms like a tensor with one upper index

2. Move indices from the top to the bottom or vice versa, for example

$$\sum_{ij} \varepsilon_{ijpq} \phi_k^{ij} = \psi_{pqk} \quad (2.135)$$

so $\sum_{ij} \varepsilon_{ijpq} \phi_k^{ij}$ transforms like a tensor with three lower indices.

2.4.1 Generators of $SU(N)$

It can be shown that the generators of $SU(N)$ are $N \times N$, hermitian, traceless matrices (or at least can be represented by these). In general there are $N^2 - 1$ of these (because they also need to be linearly independent of one another). For example the generators of the 3-dimensional representation of $SU(2)$ are the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.136)$$

And for the 8-dimensional representation of $SU(3)$ they are the Gell-Mann matrices

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.137)$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (2.138)$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad (2.139)$$

2.4.2 Adjoint representation of $SU(N)$

Just like with $SO(N)$ we define the adjoint representation as

$$(T^a)^{bc} = -if^{abc} \quad (2.140)$$

The b - c -th element of the generator T^a is equal to a constant time the structure constants' a - b - c -th entry.

The structure constant for the 3-dimensional representation of $SU(2)$ is $2\varepsilon_{ijk}$, so

$$[\sigma_a, \sigma_b] = \sum_c 2i\varepsilon_{abc}\sigma_c \quad (2.141)$$

An interesting fact about $SU(2)$ is that it is locally isomorphic to $SO(3)$ and that it covers $SO(3)$ doubly. Hence we can represent rotations of vectors in 3-dimensions using $SU(2)$! Let $X = x\sigma_1 + y\sigma_2 + z\sigma_3$ then, for example

$$U^\dagger X U = X', \quad U = e^{i\boldsymbol{\varphi} \cdot \boldsymbol{\sigma}/2} \quad (2.142)$$

will rotate X by an angle equal to $|\boldsymbol{\varphi}|$ about the axis of rotation $\hat{\boldsymbol{\varphi}}$. When using the generators for rotations, for example

$$\exp(-i\theta J_x) J_y \exp(i\theta J_x) = \cos \theta J_y + \sin \theta J_z \quad (2.143)$$

That is this expression rotates J_y about the x -axis by the angle θ .

Facts about $SU(2)$:

1. It can be described only using symmetric tensors with upper indices
2. Its representations have dimension $2j + 1$ for $(2j) \in \mathbb{Z}$
3. All of its representations are pseudoreal.

2.5 $SU(3)$

Subatomic particles' charge is described using two quantum numbers, their isospin, I_3 and their hypercharge Y . Their charge is then

$$Q = I_3 + \frac{1}{2}Y \quad (2.144)$$

I_3 and Y are the quantum numbers that describe these particles, as they are described by $SU(3)$, which has two commuting operators.

2.5.1 *Dimensions of $SU(3)$ tensors*

In contrast to $SU(2)$ there can both be upper and lower indices in $SU(3)$ (ε has three indices, hence we cannot move everything up to the upper index). Therefore it is important that we distinguish between upper and lower indices.

Let us denote *symmetric traceless tensors with \mathbf{m} upper indices and \mathbf{n} lower indices* by (\mathbf{m}, \mathbf{n})

We have previously shown that a symmetric tensor with n indices has dimension $\frac{1}{2}(n+1)(n+2)$. Additionally we have shown that the dimension of a product is the product of dimensions, hence if we have a symmetric tensor with m upper indices and n lower indices we would expect to get $\frac{1}{4}(m+1)(n+1)(m+2)(n+2)$, but we need to remove the trace, which is a tensor with $(m-1)$ upper indices and $(n-1)$ lower indices, hence

$$\mathcal{D}(m, n) = \frac{1}{4}(m+1)(n+1)(m+2)(n+2) - \frac{1}{4}mn(m+1)(n+1) = \frac{1}{2}(m+1)(n+1)(m+n+2) \quad (2.145)$$

Here is the table from A. Zee depicting the dimensions of some of the lowest irreducible representations

(m, n)	\mathcal{D}
(1, 0)	3
(1, 1)	8
(2, 0)	6
(2, 1)	15
(2, 2)	27
(3, 0)	10

Table 2.1: Dimension of a few irreducible representations of $SU(3)$. m is the number of upper indices and n is the number of lower indices of the symmetric traceless tensor.

2.5.2 *Multiplying irreducible representations of $SU(3)$*

Let us multiply two irreducible representations (m, n) and (m', n') . How do these decompose? Let's begin with $(1, 0)$ and $(0, 1)$, that is φ^i and ψ_j . When we take the tensor product of these we

get a tensor T_j^i , which we symmetrise and remove the trace and get $(1, 1)$. And what is left is the trace, which is a scalar $(0, 0)$, thus we conclude

$$(1, 0) \otimes (0, 1) = (1, 1) \oplus (0, 0), \quad (m, n, m' \text{ and } n') \quad (2.146)$$

As before we can write this as

$$3 \otimes 3^* = 8 \oplus 1, \quad (\text{dimensions}) \quad (2.147)$$

where we differentiate between the two 3-dimensional representations, as the one has one upper index, and the other one lower index (they are conjugate to each other). A more challenging example is $(1, 1)$ and $(1, 1)$, which decompose into

$$(1, 1) \otimes (1, 1) \quad (m, n, m' \text{ and } n') \quad (2.148)$$

$$= (2, 2) \oplus (3, 0) \oplus (0, 3) \oplus (1, 1) \oplus (1, 1) \oplus (0, 0) \quad (2.149)$$

and once again

$$8 \otimes 8 = 28 \oplus 10 \oplus 10^* \oplus 8 \oplus 8 \oplus 1, \quad (\text{dimensions}) \quad (2.150)$$

By repeatedly symmetrising and subtracting traces we can do this for arbitrary (m, n) and (m', n') . But first let us introduce new notation: $(m + m', n + n') = (m, n; m', n')$, so $(m, n; m', n')$ is a traceless symmetric tensor with $(m + m')$ upper indices and $(n + n')$ lower indices. Then

$$(m, n) \otimes (m', n') = (m, n; m', n') \oplus (m - 1, n; m', n' - 1) \oplus (m, n - 1; m' - 1, n') \quad (2.151)$$

$$\oplus (m - 1, n - 1; m' - 1, n' - 1) \oplus (m, n - 2; m' - 2, n') \oplus \dots \parallel \quad (2.152)$$

the $\dots \parallel$ indicates that the process continues, but that it may terminate before $(0, 0)$, (it could for example terminate with $(1, 0)$).

2.6 WEIGHTS AND ROOTS - SU(3)

The Gell-Mann matrices are central in the 8-dimensional representation of $SU(3)$, which is an important representation in particle physics. Here I have written them out again:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.153)$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (2.154)$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad (2.155)$$

Now note:

1. λ_3 and λ_8 are diagonal – they commute
2. λ_1, λ_2 and λ_3 form a $SU(2)$ subalgebra

These generators satisfy the commutation relation that defines the Lie Algebra

$$[T^a, T^b] = \sum_c if^{abc} T^c \quad (2.156)$$

for (the 8-dimensional representation of) $SU(3)$ the structure constants are

$$f^{123} = 1 \quad (2.157)$$

$$f^{147} = -f^{156} = f^{246} = f^{257} = f^{345} = -f^{367} = \frac{1}{2} \quad (2.158)$$

$$f^{458} = f^{678} = \frac{\sqrt{3}}{2} \quad (2.159)$$

and all other values *that are not related to these through permutations of indices* are zero. The following discussion will be about the generators T^a in general (not specific to any representation). We see (from the structure constants) that $[T^3, T^8] = 0$, hence they can be diagonalised simultaneously. The remaining six generators are related to lowering and raising operators, as we saw for $SO(3)$:

$$I_{\pm} = T_1 \pm iT_2 \quad (2.160)$$

$$U_{\pm} = T_6 \pm iT_7 \quad (2.161)$$

$$V_{\pm} = T_4 \pm iT_5 \quad (2.162)$$

$$I_3 = T_3 \quad (2.163)$$

$$Y = \frac{2}{\sqrt{3}}T_8 \quad (2.164)$$

The odd normalisation of Y has historical reasons. Remember that T_1, T_2 and T_3 formed a $SU(2)$ subalgebra? Well that means that

$$[I_3, I_{\pm}] = \pm I_{\pm}, \quad [I_+, I_-] = 2I_3 \quad (2.165)$$

Here are the remaining commutation relations

$$[I_3, U_{\pm}] = \mp \frac{1}{2}U_{\pm}, \quad [I_3, V_{\pm}] = \pm \frac{1}{2}V_{\pm} \quad (2.166)$$

$$[Y, U_{\pm}] = \pm U_{\pm}, \quad [Y, V_{\pm}] = \pm V_{\pm} \quad (2.167)$$

$$[Y, I_{\pm}] = 0, \quad [I_+, V_-] = -U_- \quad (2.168)$$

$$[I_+, U_+] = V_+ \quad [U_+, V_-] = I_- \quad (2.169)$$

$$[I_+, V_+] = 0 \quad [I_+, U_-] = 0 \quad (2.170)$$

$$[U_+, U_-] = \frac{3}{2}Y - I_3 \equiv U_3 \quad (2.171)$$

$$[V_+, V_-] = \frac{3}{2}Y + I_3 \equiv V_3 \quad (2.172)$$

$$[U_+, V_+] = 0 \quad (2.173)$$

2.6.1 Root vectors

We are going to construct a lattice that represents all the states in the system. On the x -axis we will have i_3 (the value that I_3 spits out) and on the y -axis we will have i_8 (the value that I_8 spits out).

Thus we call our states $|i_3, i_8\rangle$, and they are eigenstates to both I_3 and I_8 . The **root vectors** are in which direction the different raising and lowering operators take us:

$$I_{\pm} \leftrightarrow \begin{pmatrix} \pm 1 \\ 0 \end{pmatrix} \quad (2.174)$$

$$U_{\pm} \leftrightarrow \begin{pmatrix} \mp \frac{1}{2} \\ \pm 1 \end{pmatrix} \quad (2.175)$$

$$V_{\pm} \leftrightarrow \begin{pmatrix} \pm \frac{1}{2} \\ \pm 1 \end{pmatrix} \quad (2.176)$$

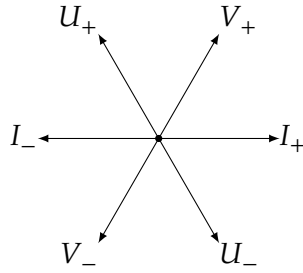


Figure 2.1: Root vectors for SU(3)

Now that we know in which direction we can move, let us find out where we can start and stop, such that we will have found all the states. The states are represented by **weights** (vertices). But before we look at that let us define two terms:

1. Positive roots: roots whose first nonzero value is positive
2. Simple roots: a positive root that cannot be written as a sum of two positive roots *with positive coefficients*.

Our positive roots are V_+ , I_+ and U_- and the simple roots are V_+ and U_- .

Next: weight diagrams are specific to which representation we are looking at! Remember we write our representation as (m, n) where m is the number of upper indices and n is the number of lower indices. Now it turns out that (m, n) also is the number of times we can use V_- and U_+ respectively on the **highest weight**. For example $(1, 0)$, which is the defining 3-dimensional representation, there we can use V_- once and U_+ no times. But how do we find the highest weight? You find it by using

the highest weight state is the state that is farthest along the line $i_3 = i_8$

$$\frac{\alpha_i \cdot \mu^{HW}}{\|\alpha_i\|^2} = \frac{q_i}{2} \quad (2.177)$$

where α_i is the i -th simple root and q_i is the i -th entry in our "vector" (m, n) . Hence for $(1, 0)$

$$\frac{1}{2} (\mu_1^{HW} + \sqrt{3} \mu_2^{HW}) = \frac{1}{2} \quad (2.178)$$

$$\frac{1}{2} (\mu_1^{HW} - \sqrt{3} \mu_2^{HW}) = 0 \quad (2.179)$$

Hence $\mu^{HW} = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}} \right)$. Now we use two important facts

1. There is a $SU(2)$ subalgebra along the i_3 -axis
2. We can use V_- and U_+ m and n times, respectively

Thus our weight diagram for $(1, 0)$ is as follows

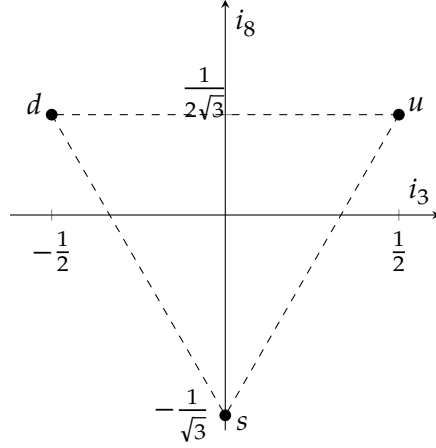


Figure 2.2: Weight diagram for the 3-dimensional representation of $SU(3)$, $(1, 0)$. u is the up quark, d the down quark and s the strange quark

Here comes another important fact about weight diagrams. To get from a representation to its conjugate you merely mirror it along the i_3 axis. This can be seen by looking at the commutation relations

$$[D(T^a), D(T^b)] = \sum_c if^{abc} D(T^c) \quad (2.180)$$

note that $[A, B]^* = (AB - BA)^* = A^*B^* - B^*A^* = [A^*, B^*]$

Here I am representing the T 's by the matrices, D . Now let us take the complex conjugate

$$[D^*(T^a), D^*(T^b)] = - \sum_c if^{abc} D^*(T^c) \quad (2.181)$$

which we can write as

$$[-D^*(T^a), -D^*(T^b)] = \sum_c if^{abc} (-D^*(T^c)) \quad (2.182)$$

This changes the signs of *both* i_3 and i_8 , however, note that the weight diagrams always are symmetric about the i_8 -axis, therefore we effectively only mirror about the i_3 axis. You can find further weight diagrams in the appendix.

2.7 WEIGHTS AND ROOTS -- GENERAL LIE ALGEBRAS

First we must introduce new notation. Let H^i be the generators that commute, and let us call the remaining generators E^i . In the case of (the 3-dimensional representation of) $SU(3)$ we have that

$$H^1 = \frac{1}{\sqrt{2}} \text{diag}(1, -1, 0) \quad (2.183)$$

$$H^2 = \frac{1}{\sqrt{6}} \text{diag}(1, 1, -2) \quad (2.184)$$

$$(2.185)$$

That is to say the the H 's are diagonal matrices with the given values on the diagonal. The weights are found by reading off vertically, that is

$$w^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}, w^2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}, w^3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -\frac{2}{\sqrt{3}} \end{pmatrix} \quad (2.186)$$

In our case here we only have three weights, so the six roots are easy to calculate (they are all the possible vectors that join the three points):

$$\alpha^1 = -\alpha^4 = w^1 - w^2 = \sqrt{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (2.187)$$

$$\alpha^2 = -\alpha^5 = w^2 - w^3 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \quad (2.188)$$

$$\alpha^3 = -\alpha^6 = w^1 - w^3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix} \quad (2.189)$$

Corresponding to $\sqrt{2}I_{\pm}$, $\sqrt{2}U_{\pm}$ and $\sqrt{2}V_{\pm}$ respectively. $SO(4)$ can be done similarly, and we end up with a square weight diagram, that can be seen in the appendix

2.7.1 Generally for $SO(2l)$

The commuting generators are

$$H^1 = \text{diag}(1, -1, 0, 0, \dots, 0, 0) \quad (2.190)$$

$$H^2 = \text{diag}(0, 0, 1, -1, 0, 0, \dots, 0, 0) \quad (2.191)$$

$$\vdots \quad (2.192)$$

$$H^{\ell} = \text{diag}(0, 0, \dots, 0, 0, 1, -1) \quad (2.193)$$

so the weights are

$$w^1 = (1, 0, \dots, 0) \quad (2.194)$$

$$w^2 = (-1, 0, \dots, 0) \quad (2.195)$$

$$w^3 = (0, 1, 0, \dots, 0) \quad (2.196)$$

$$w^{2\ell-1} = (0, \dots, 0, 1) \quad (2.197)$$

$$w^{2\ell} = (0, \dots, 0, -1) \quad (2.198)$$

$$(2.199)$$

($\omega^i \in \mathbb{R}^{\ell}$), thus we can write the weights as $\pm e^i$ where e^i is the i -th unit vector in \mathbb{R}^{ℓ} . Thus the roots are given by

$$\pm e^i \pm e^j \text{ (signs uncorrelated) } (i < j) \quad (2.200)$$

there are $4\ell(\ell-1)/2$ of these. We define the positive roots as $(e^i \pm e^j)$ again with $i < j$, which makes the simple roots:

$$e^{i-1} - e^i, e^{\ell-1} + e^{\ell}, \quad i \in \{1, 2, \dots, \ell\} \quad (2.201)$$

2.7.2 And now $SO(2l + 1)$

And all the other matrices
get a column and row of
zeros added to the end

$SO(2l + 1)$ is very similar to $SO(2l)$ but we have an additional matrix

$$H^{\ell+1} = \text{diag}(0, 0, \dots, 0) \quad (2.202)$$

(remember they have to be traceless). And hence another weight

$$w^{\ell+1} = (0, 0, \dots, 0) \quad (2.203)$$

the roots for $SO(2l + 1)$ are given by

$$\pm e^i \pm e^j \text{ (signs uncorrelated) } (i < j), \pm e^i \quad (2.204)$$

and hence the positive roots are $e^i \pm e^j$ and e^i , still with $i < j$, and the simple roots are

$$e^{i-1} - e^i, \quad e^\ell \quad (2.205)$$

2.7.3 The roots of $SU(N)$

There are $\ell = N - 1$ traceless matrices that commute with one another, these are

$$H^1 = \frac{1}{\sqrt{2}} \text{diag}(1, -1, 0, \dots, 0) \quad (2.206)$$

$$H^2 = \frac{1}{\sqrt{6}} \text{diag}(1, 1, -2, 0, \dots, 0) \quad (2.207)$$

$$\vdots \quad (2.208)$$

$$H^i = \frac{1}{\sqrt{i(i+1)}} \text{diag}(1, 1, 1, \dots, 1, -i, 0, \dots, 0) \quad (2.209)$$

$$\vdots \quad (2.210)$$

$$H^{\ell-1} = \frac{1}{\sqrt{(\ell-1)\ell}} \text{diag}(1, 1, \dots, 1, 1, 1, -(\ell-1), 0) \quad (2.211)$$

$$H^\ell = \frac{1}{\sqrt{\ell(\ell+1)}} \text{diag}(1, 1, \dots, 1, 1, 1, 1, -\ell) \quad (2.212)$$

which makes the weights

$$w^1 = \sqrt{2} \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}, \dots, \frac{1}{\sqrt{2m(m+1)}}, \dots, \frac{1}{\sqrt{2\ell(\ell+1)}} \right) \quad (2.213)$$

$$w^2 = \sqrt{2} \left(\frac{-1}{2}, \frac{1}{2\sqrt{3}}, \dots, \frac{1}{\sqrt{2m(m+1)}}, \dots, \frac{1}{\sqrt{2\ell(\ell+1)}} \right) \quad (2.214)$$

$$w^3 = \sqrt{2} \left(0, -\frac{1}{\sqrt{3}}, \dots, \frac{1}{\sqrt{2m(m+1)}}, \dots, \frac{1}{\sqrt{2\ell(\ell+1)}} \right) \quad (2.215)$$

$$\vdots \quad (2.216)$$

$$w^{m+1} = \sqrt{2} \left(0, 0, \dots, 0, \frac{-m}{\sqrt{2m(m+1)}}, \dots, \frac{1}{\sqrt{2\ell(\ell+1)}} \right) \quad (2.217)$$

$$\vdots \quad (2.218)$$

$$w^{\ell+1} = \sqrt{2} (0, 0, 0, 0, 0, \dots, \frac{-\ell}{\sqrt{2\ell(\ell+1)}}) \quad (2.219)$$

That's chaos, so let's look at $SU(4)$. $\ell = N - 1$, thus we expect 3 commuting operators:

$$H^1 = \frac{1}{\sqrt{2}} \text{diag}(1, -1, 0, 0) \quad (2.220)$$

$$H^2 = \frac{1}{\sqrt{6}} \text{diag}(1, 1, -2, 0) \quad (2.221)$$

$$H^3 = \frac{1}{2\sqrt{3}} \text{diag}(1, 1, 1, -3) \quad (2.222)$$

and the weights are

$$w^1 = \frac{1}{\sqrt{2}} \left(1, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{6}} \right) \quad (2.223)$$

$$w^2 = \frac{1}{\sqrt{2}} \left(-1, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{6}} \right) \quad (2.224)$$

$$w^3 = \frac{1}{\sqrt{2}} \left(0, -\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{6}} \right) \quad (2.225)$$

$$w^4 = \frac{1}{\sqrt{2}} \left(0, 0, -\sqrt{\frac{3}{2}} \right) \quad (2.226)$$

There are six roots, of which 3 are simple:

$$\alpha^1 = \sqrt{2}(1, 0, 0) \quad (2.227)$$

$$\alpha^2 = \frac{1}{\sqrt{2}}(-1, \sqrt{3}, 0) \quad (2.228)$$

$$\alpha^3 = \frac{1}{\sqrt{3}}(0, -\sqrt{2}, 2) \quad (2.229)$$

$$(2.230)$$

Generally for $SU(\ell + 1)$ the ℓ simple roots are

$$\alpha^i = e^i - e^{i+1}, \quad i \in \{1, 2, \dots, \ell\} \quad (2.231)$$

2.7.4 Symplectic Algebras ($Sp(2l)$)

The commuting generators are

$$H^i = u^i \otimes \sigma_3 = \left(\begin{array}{c|c} u^i & 0 \\ \hline 0 & -u^i \end{array} \right) \quad (2.232)$$

where $(u^i)_{mn} = \delta^{im} \delta^{jn}$, i.e u^i is an $\ell \times \ell$ matrix with a single entry: a one in the i -th column and i -th row. For example for $Sp(4)$:

$$H^1 = \text{diag}(1, 0, -1, 0), \quad H^2 = \text{diag}(0, 1, 0, -1) \quad (2.233)$$

which makes that weights

$$w^1 = (1, 0), \quad w^2 = (0, 1), \quad w^3 = (-1, 0), \quad w^4 = (0, -1) \quad (2.234)$$

and the positive roots are

$$\alpha^1 = 2e^1, \quad \alpha^2 = e^1 + e^2, \quad \alpha^3 = 2e^2, \quad \alpha^4 = e^1 - e^2 \quad (2.235)$$

To summarise:

	# of generators	roots	simple roots
$SU(\ell)$	$\ell^2 - 1$	$e^i - e^j$	$e^i - e^{i+1}$
$SO(2\ell + 1)$	$\ell(2\ell + 1)$	$\pm e^i \pm e^j, \pm e^i$	$e^{i-1} - e^i, e^\ell$
$Sp(2\ell)$	$\ell(2\ell + 1)$	$\pm e^i \pm e^j, \pm 2e^i$	$e^{i-1} - e^i, 2e^\ell$
$SO(2\ell)$	$\ell(2\ell - 1)$	$\pm e^i \pm e^j$	$e^{i-1} - e^i, e^{\ell-1} + e^\ell$

Table 2.2

2.8 CLASSIFICATION OF LIE ALGEBRAS

Remember a Lie Algebra with n generators is defined by

$$[T^a, T^b] = \sum_c if_c^{ab} X^c \quad (2.236)$$

f_c^{ab} is antisymmetric in ab , but it does not make sense to swap for example a and c around. We introduce now the **Cartan-Killing metric**:

$$g^{ab} \equiv \text{tr } T^a T^b = - \sum_{c,d} f_d^{ac} f_c^{bd} \quad (2.237)$$

A. Zee says that if g^{ab} is real and symmetric we can relate it to, by a similarity transformation, δ^{ab} , so that the space in question is effectively Euclidian flat space.

In physics we restrict ourselves to Lie Algebras for which g^{ab} is invertible

$$\sum_b g^{ab} g_{bc} = \delta_c^a \quad (2.238)$$

2.8.1 Cartan Subalgebra

As before we denote the commuting generators by H^i and the remaining generators by E , i.e we have

$$[H^i, H^j] = 0 \quad (2.239)$$

At this stage A. Zee starts using T^i for the matrices that represent H^i , but I will use H^i because that's far less confusing, but beware that H^i is the generator and T_i is the matrix representation thereof. We know that these are diagonal, let us call the a -th diagonal element $\beta^i(a)$

$$(H^i)_b^a = \begin{pmatrix} \beta^i(1) & 0 & 0 & 0 \\ 0 & \beta^i(2) & 0 & 0 \\ 0 & 0 & \beta^i(3) & 0 \\ 0 & 0 & 0 & \beta^i(4) \end{pmatrix} = -\beta^i(a) \delta_b^a \quad (2.240)$$

This means that

$$[H^i, X^a] = \sum_b \beta^i(a) \delta_b^a X^b = \beta^i(a) X^a \quad (2.241)$$

X^a is one of the non-commuting matrices, which we henceforth will denote by E_β . Additionally we will define the root vector $\beta(a)$:

$$\beta(a) \equiv (\beta^1(a), \beta^2(a), \dots, \beta^\ell(a)) \quad (2.242)$$

Remember before, when we took the i -diagonal entry in the H matrices and called that the root, that is exactly what this definition says. We have that

$$[H^i, E_\beta] = \beta^i E_\beta \leftrightarrow [H^i, E_\beta^\dagger] = -\beta^i E_\beta^\dagger \quad (2.243)$$

where you get the right part by taking the complex conjugate. Note that this means that

$$E_\beta^\dagger \equiv E_{-\beta} \quad (2.244)$$

A cool fact about roots is that if E_α and E_β are roots, then so is $[E_\alpha, E_\beta]$:

$$[H^i, [E_\alpha, E_\beta]] = -[E_\alpha, [E_\beta, H^i]] - [E_\beta, [H^i, E_\alpha]] \quad (2.245)$$

$$= [E_\alpha, \beta^i E_\beta] - [E_\beta, \alpha^i E_\alpha] \quad (2.246)$$

$$= (\alpha + \beta)^i [E_\alpha, E_\beta] \quad (2.247)$$

Thus $[E_\alpha, E_\beta]$ is associated with the root $(\alpha + \beta)$. However iff $(\alpha + \beta)$ is not a root, then $[E_\alpha, E_\beta] = 0$. Thus 2α is not a root (all operators commute with themselves). The following is true about the commutators

$$[H^i, H^j] = 0 \quad (2.248)$$

$$[H^i, E_\alpha] = \alpha^i E_\alpha \quad (2.249)$$

$$[E_\alpha, E_\beta] = N_{\alpha, \beta} E_{\alpha + \beta} \quad (2.250)$$

$$[E_\alpha, E_{-\alpha}] = \alpha_i H^i \quad (2.251)$$

Where $N_{\alpha, \beta}$ is some constant and $\sum_i \alpha_i g^{ij} = \alpha^j$. Let's suppose that we start at some β and the use E_α on it, this will move is in the direction of α , but not indefinitely, because the number of weights is finite, hence the ladder must terminate at some point. Let p be the number of times we can use E_α and q the number of times we can use $E_{-\alpha}$. This allows us to write an inductive relation for $N_{\alpha, \beta}$, which leads to

$$2 \frac{\langle \alpha | \beta \rangle}{\langle \alpha | \alpha \rangle} = q - p = n \in \mathbb{Z} \quad (2.252)$$

similarly

$$2 \frac{\langle \alpha | \beta \rangle}{\langle \beta | \beta \rangle} = q' - p' = m \in \mathbb{Z} \quad (2.253)$$

Hence

$$\cos^2 \theta_{\alpha\beta} = \frac{mn}{4} \implies 0 \leq mn \leq 4 \quad (2.254)$$

where $\theta_{\alpha\beta}$ is the angle between α and β . Additionally this implies that

$$\rho_{\alpha\beta} = \frac{\langle \alpha | \alpha \rangle}{\langle \beta | \beta \rangle} = \frac{m}{n} \in \mathbb{Q} \quad (2.255)$$

This is remnant of the \hat{a} and \hat{a}^\dagger for the quantum mechanical harmonic oscillator, and that their action is the opposite of one another (the one is the creation and the other the annihilation operator)

This leads to the **Cartan Killing classification of Lie Algebras**: We can determine the angles between the roots, which puts all Lie Algebras in one of four groups, this is summarised in the following table

m	n	$\frac{\langle \alpha \alpha \rangle}{\langle \beta \beta \rangle}$	$\cos^2 \theta_{\alpha\beta}$	$\theta_{\alpha\beta}$
1	1	1	$\frac{1}{4}$	60°
2	1	2	$\frac{1}{2}$	45°
3	1	3	$\frac{1}{4}$	30°
2	2	1	1	0°

Table 2.3: Summary of the geometry of root diagrams.

Note that with this table we can show that:

There are at most two different lengths of roots in a root diagram.

A very useful tool when determining root diagrams is **Weyl reflection**: Given two roots α and β we define

$$\beta' = \beta - 2 \frac{\langle \alpha | \beta \rangle}{\langle \alpha | \alpha \rangle} \alpha = \beta + (p - q)\alpha \quad (2.256)$$

if β is a root, then so is β' , this helps us draw root diagrams much quicker, as it gives us additional roots, where we know the direction and the magnitude. This can be interpreted geometrically as: mirroring β through the hyperplane that is orthogonal α gives you β' .

You can find all possible root diagrams for rank-2 Lie Algebras in the appendix.

Note that the angle between two *simple* roots has to be obtuse or right:

$$\theta_{\alpha\beta} \in \left[\frac{\pi}{2}, \pi \right] \quad (2.257)$$

2.9 DYNKIN DIAGRAMS

Every simple root is assigned either a filled circle, \bullet , or an open circle, \circ . Each circle is connected with zero, one, two or three lines. Here is how we determine which:

1. If the simple root is one of the *short roots*, the circle is filled. If it is one of the *long roots* it is open.
2. Number of lines between circles:
 - a) Zero if $\theta_{\alpha\beta} = 90^\circ$
 - b) One if $\theta_{\alpha\beta} = 120^\circ$
 - c) Two if $\theta_{\alpha\beta} = 135^\circ$

d) Three if $\theta_{\alpha\beta} = 150^\circ$

The number of lines between two circles corresponding to the simple roots α and β respectively is given by

$$\mathcal{N}_L(\alpha, \beta) = \left(2 \frac{\langle \alpha | \beta \rangle}{\langle \alpha | \alpha \rangle}\right) \left(2 \frac{\langle \alpha | \beta \rangle}{\langle \beta | \beta \rangle}\right) = pp' \quad (2.258)$$

(Note that q and q' are zero because the roots are simple) A few Dynkin Diagrams are shown in the appendix.

Here are a few rules for Dynkin Diagrams:

1. Cutting: If you cut any line between two circles the resulting diagrams must be legal Dynkin Diagrams
2. No circle can have more than three lines
3. No loops

2.9.1 Cartan Matrix

We define the **Cartan Matrix** as

$$A_{ij} \equiv 2 \frac{\langle \alpha_i | \alpha_j \rangle}{\langle \alpha_i | \alpha_i \rangle} \quad (2.259)$$

2.10 THINGS NO ONE CAN REMEMBER (APPENDIX)

2.10.1 Integers modulo n — \mathbb{Z}_n

The integers form a finite group under addition modulo n . For example

$$\mathbb{Z}_4 = \{0, 1, 2, 3\} \quad (2.260)$$

because all numbers are either 0, 1, 2 or 3 mod 4. For example

$$2 + 3 \equiv 1 \pmod{4} \quad (2.261)$$

\mathbb{Z}_n is isomorphic to the n roots of unity under multiplication. For example \mathbb{Z}_4 is isomorphic to

$$\{1, i, -1, i\} \quad (2.262)$$

under multiplication (note that these are all the fourth roots of 1)

2.10.2 *Multiplication table*

	I	A	B	C
I	I	A	B	C
A	A	A^2	AB	AC
B	B	BA	B^2	BC
C	C	CA	CB	C^2

For groups of even order there is at least one element that squares to the identity.

There are two possible ways of filling this table out (There are two groups distinct 4-th order groups up to an isomorphism). Let us first assume that $A^2 = I$, then $\{I, A\}$ form a subgroup. Then $AB = C$ because the row with AB has A and I , and the column with AB has B , by similar argument $AC = B$. Also $AB = BA$ and $AC = CA$, due to the symmetry, which isn't always the case. Now we are left with a choice, either $B^2 = C^2 = I$ or $BC = CB = I$. These yield two distinct groups:

	I	A	B	C
I	I	A	B	C
A	A	I	C	B
B	B	C	I	A
C	C	B	A	I

Table 2.4: Multiplication table for any group that is isomorphic to $\mathbb{Z}_2 \otimes \mathbb{Z}_2$

	I	A	B	C
I	I	A	B	C
A	A	I	C	B
B	B	C	A	I
C	C	B	I	A

Table 2.5: The multiplication table for the cyclic group of order 4. B is the generator.

2.10.3 *Character table Q*

First we need to determine the number of equivalence classes. Note that

$$-i^2 = -j^2 = -k^2 = 1 \quad (2.263)$$

So let us just calculate all the possible similarity transformations

	1	-1	i	j	k
$1^{-1} \ 1$	1	-1	i	j	k
$(-1)^{-1} \ -1$	1	-1	i	j	k
$i^{-1} \ i$	1	-1	i	$-j$	$-k$
$j^{-1} \ j$	1	-1	$-i$	j	$-k$
$k^{-1} \ k$	1	-1	$-i$	$-j$	k

Table 2.6: For example the row j^{-1} j takes 1, i , j and k and calculates $j^{-1}1j$, $j^{-1}ij$, $j^{-1}jj$ and $j^{-1}kj$ respectively

Hence we see that

$$i \equiv -i, \quad j \equiv -j, \quad k \equiv -k \quad (2.264)$$

hence we have five equivalence classes, and hence five irreducible representations ($N(C) = N(R)$). Now we use that

$$\sum_r d_r^2 = N(G) \quad (2.265)$$

This says we have five numbers whose sum of squares must be equal to eight (there are eight elements in the group). There is only one solution to this

$$1 + 1 + 1 + 1 + 2^2 = 8 \quad (2.266)$$

Thus we can already fill out a lot in the character table

Q	n_c	c	1	1'	1''	1'''	2
	1	1	1	1	1	1	2
	1	-1	1	w	x	y	z
	2	i	1	α	β	γ	ϵ
	2	j	1	a	b	c	d
	2	k	1	Λ	Γ	Ξ	Σ

Table 2.7: Character table for Q , incomplete.

First off we use

$$\sum_c n_c \chi^{(r)*}(c) \chi^{(s)}(c) = N(G) \delta^{rs} \quad (2.267)$$

which, when we let $r = s$ tells us that all entries in the columns $1'$, $1''$ and $1'''$ either are 1 or -1 . Also $z = \pm 2$. Now if we use the same equation but choose $r = 1$ and $s = 2$ we see that $z = -2$, so that the sum is zero. Now let us use

$$\sum_r (\chi^{(r)}(c))^* \chi^{(r)}(c') = \frac{N(G)}{n_c} \delta^{cc'} \quad (2.268)$$

with together with what we established previously tells that $w = x = y = 1$. Now we know that the remaining entrances have to have -1 twice in each row and column (due to the orthogonality), it doesn't matter in which order we do it, because we can freely swap the columns $1'$, $1''$ and $1'''$ around. Therefore the complete character table is

Q	n_c	c	1	1'	1''	1'''	2
	1	1	1	1	1	1	2
	1	-1	1	1	1	1	-2
	2	i	1	1	-1	-1	0
	2	j	1	-1	1	-1	0
	2	k	1	-1	-1	1	0

Table 2.8: Character table for Q , complete.

2.10.4 Further Weight Diagrams

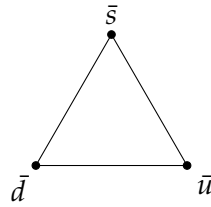


Figure 2.3: 3-dimensional representation of $SU(3)$, $(0, 1)$*

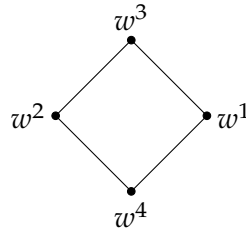


Figure 2.4: 4-dimensional representation of $SO(4)$

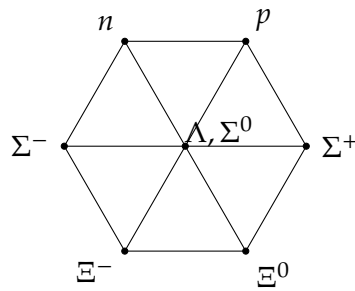


Figure 2.5: 8-dimensional representation of $SU(3)$, $(1, 1)$. Note that there are two distinct states in the middle, Λ and Σ^0

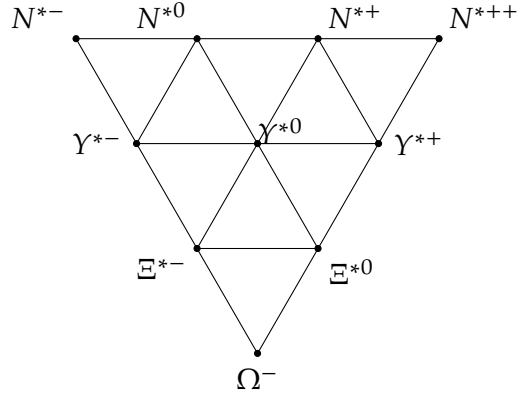


Figure 2.6: 10-dimensional representation of $SU(3)$, $(3,0)$

2.10.5 Root diagrams for rank-2 Lie Algebras

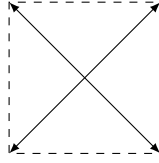


Figure 2.7: $SO(4) = D_4$, $\theta_{\alpha\beta} = 90^\circ$

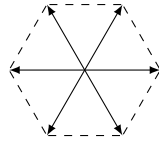


Figure 2.8: $SU(3) = A_2$, $\theta_{\alpha\beta} = 60^\circ$

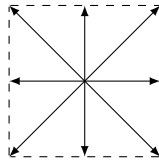


Figure 2.9: $SO(5) = B_2$, $\theta_{\alpha\beta} = 45^\circ$

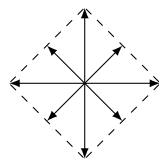


Figure 2.10: $Sp(4) = C_2$, $\theta_{\alpha\beta} = 45^\circ$

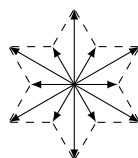


Figure 2.11: G_2 , $\theta_{\alpha\beta} = 30^\circ$

2.10.6 A Few Dynkin Diagrams

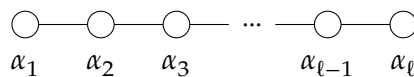


Figure 2.12: $SU(l+1)$

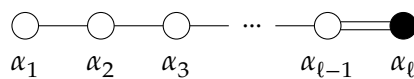


Figure 2.13: $SO(2l+1)$

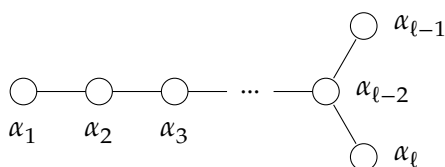


Figure 2.14: $SO(2l)$. If $l = 2$ then α_{l-2} does not exist, hence the two circles are not connected.

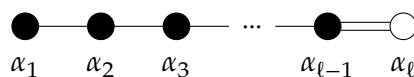


Figure 2.15: $Sp(2l)$

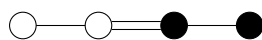


Figure 2.16: F_4

ATOMIC PHYSICS

3.1 HYDROGEN ATOM

J. Rydberg found experimentally that the spectral lines in hydrogen are given by

$$\frac{1}{\lambda} = R \left(\frac{1}{n} - \frac{1}{m} \right), \quad n, m \in \mathbb{Z}, n < m \quad (3.1)$$

The *Balmer series* is the series $n = 2, m = 3, 4, \dots$ and the lines are named using the greek alphabet, such that the first line, $\lambda = 656\text{nm}$ is the Balmer- α line (H_α). Subsequent lines have shorter and short wavelengths. Often one defines the wavenumber $\tilde{\nu}$:

$$\tilde{\nu} = \lambda^{-1} \quad (3.2)$$

as this is closely related to the energy of the transition:

$$E = hc\tilde{\nu} \quad (3.3)$$

Other series are more difficult for humans to observe (with their eyes): The *Lyman series*, which has $n = 1$ lies in the ultraviolet region, and the series with wave lengths longer than the Balmer series lie in and beyond the infra-red region.

3.1.1 Bohr's Model

Bohr assumed electrons were small masses that orbited the nucleus, setting the centripetal force equal and opposite to the Coulomb force you get:

$$\frac{m_e v^2}{r} = \frac{e^2}{4\pi\epsilon_0 r^2} \quad (3.4)$$

Now, using the classical Hamiltonian ($H = T + V$) Bohr showed that the total energy of the electron in orbit about the hydrogen nucleus would have negative energy:

$$E = -\frac{1}{2r} \frac{e^2}{4\pi\epsilon_0} \quad (3.5)$$

In order to remove the dependence on the unknown quantity (r), Bohr assumed that the angular momentum was quantised in units of \hbar :

$$m_e v r = n\hbar, \quad n \in \mathbb{Z} \quad (3.6)$$

which implies

$$r = a_n n, \quad a_0 = \frac{\hbar^2}{m_e} \frac{4\pi\epsilon_0}{e^2} \quad (3.7)$$

where a_0 is the **Bohr radius**. This gives us the following expression for the hydrogen energy:

$$E = -\frac{1}{2a_0n^2} \frac{e^2}{4\pi\epsilon_0} \quad (3.8)$$

This fits very closely to what was observed experimentally, as we can define the Rydberg constant as

$$hcR_\infty = \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m_e}{2\hbar^2} \quad (3.9)$$

The subscript R_∞ denotes that the Bohr model has assumed that the nucleus (one proton) has infinite mass – i.e the proton is stationary. In reality though the electron-proton pair orbit about their centre of mass. The classical Hamilton for the two body problem is normally solved using the reduced mass:

$$\mu = \frac{m_e M}{m_e + M} \quad (3.10)$$

where M is the mass of the nucleus. Hence for hydrogen we get a correction that is of order $m_e/M \approx \frac{1}{1836}$:

$$R_H = R_\infty \frac{M}{m_e + M} \approx R_\infty \left(1 - \frac{m_e}{M} \right) \quad (3.11)$$

This leads to a small but observable shift in energy-levels between different isotopes, known as the **isotope shift**.

3.1.2 Einstein's A and B coefficients

Consider a two-level atomic system, with populations N_1 and N_2 of the lower and upper levels respectively. These atoms are placed into an environment with a continuous radiation density $\rho(\omega)$. The excitation $|1\rangle \rightsquigarrow |2\rangle$ occurs with a rate that is proportional to $\rho(\omega_{12})$ where $\omega_{12} \equiv (E_2 - E_1)/\hbar$. The proportionality constant is B_{12} . Similarly the radiation promotes decay $|2\rangle \rightsquigarrow |1\rangle$ at a rate proportional to $\rho(\omega_{12})$, however, with proportionality constant B_{21} . Additionally there is a non-zero probability to decay spontaneously; we treat this effect qualitatively, and merely say that this occurs at a rate $\tau = \frac{1}{A_{21}}$. Hence

$$\frac{dN_2}{dt} = N_1 B_{12} \rho(\omega_{12}) - N_2 B_{21} \rho(\omega_{12}) - N_2 A_{21} \quad (3.12)$$

additionally, because the total number of atoms is constant

$$\frac{dN_2}{dt} = -\frac{dN_1}{dt} \quad (3.13)$$

Let us look at an example: Assume that $\rho(\omega) = 0$ (the radiation is switched off), and that there are some excited atoms, say $N_2(0)$ of these. This means that

$$N_2(t) = N_2(0)e^{-A_{21}t} \quad (3.14)$$

Next let us look at the rate equations within a black body. That is, let us assume that the radiation density is given by Planck's distribution law:

$$\rho(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega/k_B T) - 1} \quad (3.15)$$

This gives us an expression for $\rho(\omega_{12})$ at equilibrium. Additionally we must use that the population of each state (divided by its degeneracy) is given by the Boltzmann factor :

$$\frac{N_2}{g_2} = \frac{N_1}{g_1} \exp\left(-\frac{\hbar\omega}{k_B T}\right) \quad (3.16)$$

From this we find

$$A_{21} = \frac{\hbar\omega^3}{\pi^2 c^3} B_{21} \quad (3.17)$$

and

$$B_{12} = \frac{g_2}{g_1} B_{21} \quad (3.18)$$

3.2 INTERACTION OF ATOMS WITH RADIATION

Suppose our time-dependent Hamiltonian can be split in to a constant H_0 term, and a perturbative time-dependent part, $H_I(t)$. Using H_0 wave functions we can express the wave-function similarly to how we would have expressed it without $H_I(t)$'s effect, however, now our coefficients are time-dependent:

$$\Psi(t) = c_1(t)|1\rangle e^{-i\omega_1 t} + c_2(t)|2\rangle e^{-i\omega_2 t}, \quad \omega_i = \frac{E_i}{\hbar} \quad (3.19)$$

We still have the normalisation condition:

$$|c_1(t)|^2 + |c_2(t)|^2 = 1 \quad (3.20)$$

Hence it would be beneficial to describe any time-dependent perturbation by solving the differential equations for $c_i(t)$.

3.2.1 Oscillating electric field

An oscillatory electric field, $\mathbf{E}(t) = \mathbf{E}_0(t) \cos(\omega t)$, produces the following perturbation:

$$H_I(t) = e\mathbf{r} \cdot \mathbf{E}_0 \cos(\omega t) \quad (3.21)$$

which gives us the following Schrödinger equation

$$\dot{c}_1 = -i\Omega \cos(\omega t) e^{-i\omega_0 t} c_2 \quad (3.22)$$

$$\dot{c}_2 = -i\Omega^* \cos(\omega t) e^{i\omega_0 t} c_1 \quad (3.23)$$

where the **Rabi frequency** is defined as

$$\Omega = \frac{e}{\hbar} \langle 1|\mathbf{r} \cdot \mathbf{E}_0|2\rangle \quad (3.24)$$

The **dipole approximation** simplifies these calculations significantly, by assuming that the amplitude of the electric field does not change on the scale of the problem (we assume $\|\mathbf{E}_0\|$ is approximately constant, which allows to take it outside the integral), hence

$$\Omega = \frac{eX_{12}\|\mathbf{E}_0\|}{\hbar}, \quad X_{12} \equiv \langle 1|\mathbf{r}|2\rangle \quad (3.25)$$

commonly we choose $\mathbf{E} \parallel \mathbf{x}$, which turns X_{12} into $\langle 1|x|2\rangle$.

3.2.2 Rotating wave approximation

The solution to Equations 3.22 and 3.23, with the initial conditions $c_1(0) = 1$ and $c_2(0) = 0$ is

$$c_1(t) = 1 \quad (3.26)$$

$$c_2(t) = \frac{\Omega^*}{2} \left(\frac{1 - e^{i(\omega_0 + \omega)t}}{\omega_0 + \omega} + \frac{1 - e^{i(\omega_0 - \omega)t}}{\omega_0 - \omega} \right) \quad (3.27)$$

Assuming the detuning $\delta \equiv \omega_0 - \omega$ is small compared to the atomic resonance, ω_0 , we can neglect the first term:

$$|c_2(t)|^2 = \left| \Omega \frac{\sin((\omega_0 - \omega)t/2)}{\omega_0 - \omega} \right|^2 = \frac{1}{4} |\Omega|^2 t^2 \frac{\sin^2 x}{x^2} \quad (3.28)$$

which is a sinc-squared in $\omega_0 - \omega$, with a width that decreases as $t \rightarrow \infty$. In the limit $t \rightarrow \infty$ it becomes a Dirac-delta.

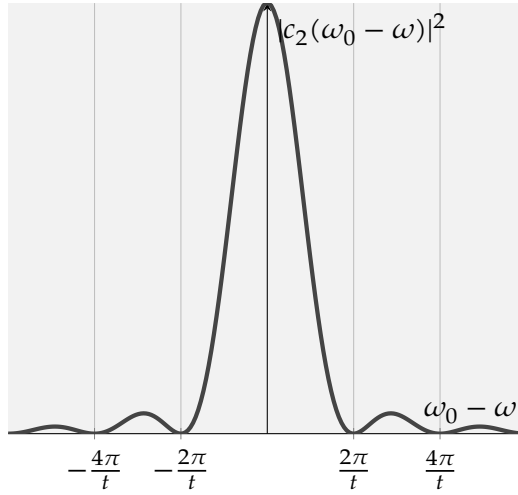


Figure 3.1: $|c_2(\omega_0 - \omega)|^2$ at a set value of t .

3.2.3 Einstein's B coefficient

By relating the radiation density $\rho(\omega)$ to the intensity, and hence the electric field, we can find an expression for $\Omega(\omega)$, which we would want to integrate over ω to get the total $c_2(t)$:

$$|c_2(t)|^2 = \frac{2e^2 |X_{12}|^2}{\epsilon_0 \hbar} \int_{\omega_0 - \Delta/2}^{\omega_0 + \Delta/2} \rho(\omega) \frac{\sin^2((\omega_0 - \omega)t/2)}{\omega_0 - \omega} d\omega \quad (3.29)$$

For large t , (in fact $t \rightarrow \infty$) the sinc function can be approximated as a Dirac delta, which spits out $\rho(\omega_0)$, giving us a transition rate of

$$R_{12} = \frac{|c_2(t)|^2}{t} = \frac{\pi e^2 |X_{12}|^2}{\epsilon_0 \hbar^2} \rho(\omega_0) \quad (3.30)$$

and hence

$$B_{12} = \frac{\pi e^2 \|\mathbf{D}_{12}\|^2}{3\epsilon_0 \hbar^2} \quad (3.31)$$

where we have used that $|X_{12}| \rightsquigarrow \|\mathbf{D}_{12}\|/3$ and

$$\mathbf{D}_{12} = \langle 1|\mathbf{r}|2\rangle \quad (3.32)$$

Our relation between A_{21} and B_{12} tells us that:

$$A_{21} = \frac{g_1}{g_2} \frac{4\alpha}{3c^2} \omega^3 |D_{12}|^2 \quad (3.33)$$

3.3 INTERACTION WITH MONOCHROMATIC RADIATION

Instead of treating the problem of the interaction between an two-level system and monochromatic radiation perturbatively, we can solve the full Schrödinger equation. This means we can allow the electric field to be strong. Initially we assumed that most of the population stayed in the ground state, i.e that $|c_1| \gg |c_2|$, however, now we will not make that approximation. Writing the Schrödinger equation:

$$\frac{dc_1}{dt} = -ic_2 \left(e^{i(\omega_0 - \omega)t} + e^{-i(\omega_0 + \omega)t} \right) \frac{\Omega}{2} \quad (3.34)$$

If ω is "close" to ω_0 the sum $\omega_0 + \omega$ will oscillate far quicker than $\omega_0 - \omega$, so the term with the sum will average out during one period of $\frac{1}{\omega_0 - \omega}$. Therefore we can ignore its effect:

$$\frac{dc_1}{dt} = -ic_2 e^{i(\omega - \omega_0)t} \frac{\Omega}{2} \quad (3.35)$$

$$\frac{dc_2}{dt} = -ic_1 e^{-i(\omega - \omega_0)t} \frac{\Omega^*}{2} \quad (3.36)$$

These coupled first-order differential equations can be decoupled to a second-order differential equation

$$\frac{d^2 c_2}{dt^2} + i(\omega - \omega_0) \frac{dc_2}{dt} + \left| \frac{\omega}{2} \right|^2 c_2 = 0 \quad (3.37)$$

Whose solution, assuming that $c_1(0) = 1$ and $c_2(0) = 0$ can be written as

$$|c_2(t)|^2 = \frac{\Omega^2}{W^2} \sin^2 \left(\frac{Wt}{2} \right), \quad W^2 = \Omega^2 + \delta^2 \quad (3.38)$$

When the detuning is zero this simplifies to

$$|c_2(t)|^2 = \sin^2 \left(\frac{\Omega t}{2} \right) \quad (3.39)$$

for $\Omega t = \pi$ we get that $|c_2(\frac{\pi}{\Omega})|^2 = 1$ and hence that $|c_1(\frac{\pi}{\Omega})|^2 = 0$ – we have inverted the population. This population inversion is more general:

$$c_1|1\rangle + c_2|2\rangle \rightsquigarrow -i(c_2|1\rangle + c_1|2\rangle) \quad (3.40)$$

For $\Omega t = \frac{\pi}{2}$ this process hasn't finished completely, as $\sin^2(\frac{\pi}{2}) = \frac{1}{2}$, hence this takes us from $|1\rangle \rightsquigarrow \frac{1}{\sqrt{2}}(|1\rangle + e^{i\phi}|2\rangle)$.

3.3.1 Bloch Sphere and Bloch Vector

This simple example from before allows us to introduce the Bloch sphere and Bloch vector rather naturally. The density matrix is defined as

$$|\Psi\rangle\langle\Psi| = \begin{pmatrix} |c_1|^2 & c_1 c_2^* \\ c_1^* c_2 & |c_2|^2 \end{pmatrix} \equiv \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \quad (3.41)$$

The diagonal entries are the populations of the two levels, and the off-diagonal elements are referred to as the *coherences*. By changing basis we can write hence rewrite the following expression for the dipole moment

$$D_x = c_1 c_2^* X_{21} e^{i\omega_0 t} + c_1^* c_2 X_{12} e^{-i\omega_0 t} \quad (3.42)$$

as follows:

$$D_x = X_{12} (u \cos(\omega t) - v \sin(\omega t)) \quad (3.43)$$

This is done by going into the *rotating frame*, which just means that we can think of our change of basis as a constant rotation with respect to the intrinsic phases of the two levels. We chose the rotating frame's frequency to be equal to the detuning $\delta = \omega - \omega_0$, hence

$$\tilde{c}_1 = c_1 e^{i\delta t/2} \quad (3.44)$$

$$\tilde{c}_2 = c_2 e^{-i\delta t/2} \quad (3.45)$$

Using this transformation we define

$$\tilde{\rho}_{11} = \rho_{11}, \quad \tilde{\rho}_{22} = \rho_{22} \quad (3.46)$$

$$\tilde{\rho}_{12} = \rho_{12} e^{-i\delta t}, \quad \tilde{\rho}_{21} = \rho_{21} e^{i\delta t} \quad (3.47)$$

Which allows us to rewrite the Schrödinger equation as

$$\frac{d}{dt} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 & \delta & 0 \\ -\delta & 0 & \Omega \\ 0 & -\Omega & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} \quad (3.48)$$

where

$$\mathbf{R} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \tilde{\rho}_{12} + \tilde{\rho}_{21} \\ -i(\tilde{\rho}_{12} - \tilde{\rho}_{21}) \\ \tilde{\rho}_{11} - \tilde{\rho}_{22} \end{pmatrix} \quad (3.49)$$

The matrix, call it M , in equation 3.48 screams "antisymmetric tensor", and can in fact be written as

$$M_{ij} = \varepsilon_{ijk} W_k, \quad \mathbf{W} = \begin{pmatrix} \Omega \\ 0 \\ \delta \end{pmatrix} \quad (3.50)$$

Thus

$$\dot{\mathbf{R}}_i = \varepsilon_{ijk} R_j W_k \rightsquigarrow \dot{\mathbf{R}} = \mathbf{R} \times \mathbf{W} \quad (3.51)$$

\mathbf{R} is the Bloch vector that describes our state (for example $\mathbf{R} = \hat{\mathbf{e}}_3$ is the state $|1\rangle$, whereas $\mathbf{R} = \hat{\mathbf{e}}_1$ is the state $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$). \mathbf{R} precesses about \mathbf{W} at angular frequency W .

3.3.2 Optical Bloch Equations

By introducing the natural decay rate Γ , to the differential equations above, we modify the differential equation that describes ρ_{22} :

$$\dot{\rho}_{22} = -\Gamma\rho_{22} + \frac{\Omega}{2}v \quad (3.52)$$

which leads to the **Optical Bloch Equations**:

$$\frac{d}{dt} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\frac{\Gamma}{2} & \delta & 0 \\ -\delta & -\frac{\Gamma}{2} & \Omega \\ 0 & -\Omega & -\Gamma \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \Gamma \end{pmatrix} \quad (3.53)$$

In steady state, i.e for $t \gg \Gamma^{-1}$ the solution reads

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \frac{1}{\delta^2 + \frac{\Omega^2}{2} + \frac{\Gamma^2}{4}} \begin{pmatrix} \Omega\delta \\ \Omega\frac{\Gamma}{2} \\ \delta^2 + \frac{\Gamma^2}{4} \end{pmatrix} \quad (3.54)$$

This is not particularly intuitive, but for example, by putting this into the definitions given above you can see that $\rho_{22} = \frac{\frac{\Omega^2}{4}}{\delta^2 + \frac{\Omega^2}{2} + \frac{\Gamma^2}{4}} \rightsquigarrow \frac{1}{2}$ for increasing intensities (and hence increasing Ω)

3.3.3 Optical absorption cross-section

The intensity of light propagating through a medium with N atoms per unit volume will decrease exponentially. The attenuation can be described by

$$\frac{dI}{dz} = -\kappa(\omega)I = -N\sigma(\omega)I \quad (3.55)$$

where $\sigma(\omega)$ is the **absorption cross-section**. Using the results from the previous sections we can express this as a function of Ω , δ and Γ :

$$\sigma(\omega) = \frac{\frac{\Omega^2}{4}}{\delta^2 + \frac{\Gamma^2}{4}} \frac{A_{21}\hbar\omega}{I} \quad (3.56)$$

Note that Ω^2 is a function of I , therefore the differential equation still has an exponential solution. Additionally note that omission of Ω^2 in the denominator, which must be because we assume that the field is sufficiently weak. We can rewrite this as

$$\sigma(\omega) = \frac{3\pi^2 c^2}{\omega_0^2} A_{21} g_H(\omega), \quad g_H(\omega) = \frac{1}{2\pi} \frac{\Gamma}{\delta^2 + \frac{\Gamma^2}{4}} \quad (3.57)$$

where the subscript, H , in the distribution functions is to show that this line width is *homogeneous* – it affects all atoms in the sample in the same way. The pre-factor of 3 can range from 0 to 3, depending on how the incident light is polarised. For unpolarised light, propagating through a two-level system, with degeneracies on both levels, the cross-section becomes

$$\sigma(\omega) = \frac{g_2}{g_1} \frac{\pi^2 c^2}{\omega_0^2} A_{21} g_H(\omega) \quad (3.58)$$

This broadening is often referred to as **natural line-width**.

3.3.4 Peak radiative broadening

For $\delta = 0$ we reach the largest cross-section:

$$\sigma(\omega_0) = \frac{6\pi^2 c^2}{\omega_0^2} \frac{A_{21}}{\Gamma} \quad (3.59)$$

In a two-level system $\Gamma = A_{21}$, because the only possible decay is from $|2\rangle \rightsquigarrow |1\rangle$. Hence

$$\sigma(\omega_0) \approx \frac{\lambda_0^2}{2}, \quad \lambda_0 \equiv \frac{2\pi c}{\omega_0} \quad (3.60)$$

This wavelength is often in the visible region, for example $\lambda_0 = 589\text{nm}$ for sodium, therefore the cross-section is often *huge* compared to the size of atoms, close to resonance. However, the cross-section decreases rapidly, off-resonance. For the example above, when we use light with wavelength $\lambda = 600\text{nm}$ the cross section is $\sigma(\omega) \sim 10^{-25}\text{m}^2$ – which is small, even compared to the Bohr radius squared.

3.3.5 Saturation Intensity

The population difference from before, can now be written as

$$w = \frac{1}{1 + \frac{I}{I_s(\omega)}} \quad (3.61)$$

where the saturation intensity is defined as

$$I_s(\omega) = \frac{\hbar\omega A_{21}}{2\sigma(\omega)} \quad (3.62)$$

which reaches its minimum at resonance – we define

$$I_{\text{sat}} \equiv \min \{I_s(\omega)\} = I_s(\omega_0) = \frac{\pi}{3} \frac{\hbar c}{\lambda_0^3 \tau}, \quad \tau = \Gamma^{-1} \quad (3.63)$$

where λ_0 is the aforementioned transition wavelength. From Equation 9.82 we recognise that

$$\frac{I}{I_{\text{sat}}} = \frac{2\Omega^2}{\Gamma^2} \quad (3.64)$$

3.3.6 Power broadening

Using what we have found, we can rearrange the expression for the attenuation $\kappa(\omega)$:

$$\kappa(\omega) = N\sigma(\omega_0) \frac{\frac{\Gamma^2}{4}}{\delta^2 + \frac{\Gamma^2}{4} \left(1 + \frac{I}{I_{\text{sat}}}\right)} \quad (3.65)$$

which is a Lorentzian distribution, and tells us that the line shape due to the radiation has a full-width-half-maximum (FWHM) of

$$\Delta\omega_{\text{FWHM}} = \Gamma \left(1 + \frac{I}{I_{\text{sat}}}\right)^{\frac{1}{2}} \quad (3.66)$$

3.4 TRANSITIONS

When looking at dipole transitions, one will have to calculate the dipole moment, as in, for example Equation 3.32. In order to do this there are some selection rules that help get an overview of which dipole-transitions are legal. First let us split the dipole moment into an angular and a radial integral:

$$\langle 2|\mathbf{r} \cdot \hat{\mathbf{e}}_{\text{rad}}|1\rangle = D_{12}\mathcal{J}_{\text{ang}} \quad (3.67)$$

where

$$D_{12} = \int_0^\infty R_{n_2,\ell_2} r R_{n_1,\ell_1} r^2 dr \quad (3.68)$$

$$\mathcal{J}_{\text{ang}} = \int_0^{2\pi} \int_0^\pi Y_{\ell_2,m_2}^*(\theta,\phi) \hat{\mathbf{r}} \cdot \hat{\mathbf{e}}_{\text{rad}} Y_{\ell_1,m_1} \sin\theta d\theta d\phi \quad (3.69)$$

The radial part measures the overlap in radial wave functions, hence for large differences in n_1 and n_2 this integral can be negligible. However, it is usually the radial integral that is zero, unless very strict selection rules hold true.

Generally the dot product $\hat{\mathbf{r}} \cdot \hat{\mathbf{e}}_{\text{rad}}$ can be written in terms of the $Y_{1,m}$ functions:

$$\hat{\mathbf{r}} \cdot \hat{\mathbf{e}}_{\text{rad}} \propto A_{\sigma^-} Y_{1,-1} + A_z Y_{1,0} + A_{\sigma^+} Y_{1,1} \quad (3.70)$$

Hence we have split up the radiation into π -radiation (linearly polarised light along the z -axis) and two σ -radiations (circular polarised light each rotating in the opposite direction)

3.4.1 π transitions

π -polarised light does not have an angular momentum along the z -axis, thus we would expect that π -transitions leave m unchanged:

$$\mathcal{J}_{\text{ang}}^\pi = \int_0^{2\pi} \int_0^\pi Y_{\ell_2,m_2} \cos\theta Y_{\ell_1,m_1} \sin\theta d\theta d\phi \quad (3.71)$$

using the cylindrical symmetry, we can show that

$$\mathcal{J}_{\text{ang}}^\pi = e^{i(m_1-m_2)\phi_0} \mathcal{J}_{\text{ang}}^\pi \quad \rightsquigarrow \quad \Delta m = 0 \quad (3.72)$$

as expected.

3.4.2 σ transitions

σ^+ -light rotates anticlockwise (seen from positive z , for light propagating in the z -direction), hence the angular momentum vector is pointing in the positive z -direction. This would lead us to believe that σ^+ -transitions increase m by one, and this is in fact the case. Similarly σ^- -transition decrease m by one:

$$\Delta m = \begin{cases} 1 & + - \text{ transitions} \\ -1 & - - \text{ transitions} \end{cases} \quad (3.73)$$

Generally therefore the angular integral is zero unless

$$\Delta m = 0, \pm 1 \quad (3.74)$$

Using the expression we got for $\hat{\mathbf{r}} \cdot \hat{\mathbf{e}}$ before we can see that

$$\mathcal{J}_{\text{ang}} \propto \int_0^{2\pi} \int_0^\pi Y_{\ell_2, m_2}^* Y_{1, m} Y_{\ell_1, m_1} \sin \theta d\theta d\phi, \quad m = 0, \pm 1 \quad (3.75)$$

Using the orthogonality of the angular functions (and a rule for multiplication of angular functions) we get an additional selection rule, which constrains $\Delta \ell$. In summary the selection rules are

$$\Delta \ell = \pm 1 \quad (3.76)$$

$$\Delta m = 0, \pm 1 \quad (3.77)$$

3.5 FINE STRUCTURE

The fine structure has three contributing components: Spin-orbit coupling, the Darwin term and a relativistic correction to the kinetic energy.

3.5.1 Spin-orbit coupling

The calculation of the correction due to spin-orbit calculation is normally carried out in the electron's reference frame, which is not an inertial reference frame. Therefore our calculations will be off by approximately a factor of two, due to what is known as **Thomas precession**.

In the electron's reference frame, the proton is moving, and therefore creates a magnetic field:

$$\mathbf{B} = \frac{\hbar}{m_e c^2} \left(\frac{1}{er} \frac{\partial V}{\partial r} \right) \ell \quad (3.78)$$

where $V(r)$ is the potential due to the proton and $\hbar \ell = m_e \mathbf{r} \times \mathbf{v}$ is the quantised angular momentum of the electron. The intrinsic magnetic moment of the electron will couple to this magnetic field. This magnetic moment is related to its spin:

$$\boldsymbol{\mu} = -g_s \mu_B \mathbf{s} \quad (3.79)$$

where μ_B is the Bohr magneton and $g_s \approx 2$. The interaction energy is given by

$$H_{\text{SO}} = -\boldsymbol{\mu} \cdot \mathbf{B} = g_s \mu_B \frac{\hbar}{m_e c^2} \left(\frac{1}{er} \frac{\partial V}{\partial r} \right) \mathbf{s} \cdot \ell \quad (3.80)$$

Evaluating the expectation value of this requires you to evaluate $\langle r^{-3} \rangle$ as well as $\langle \mathbf{s} \cdot \ell \rangle$, that latter of which can be done by introduction $\mathbf{j} = \mathbf{s} + \ell$, which is a conserved quantity. This gives us

$$E_{\text{SO}} = \frac{\beta}{2} (j(j+1) - s(s+1) - \ell(\ell+1)) \quad (3.81)$$

where the spin-orbit constant is

$$\beta = \frac{\hbar^2}{2m_e^2 c^2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{(na_0^3) \ell (\ell + \frac{1}{2}) (\ell + 1)} \quad (3.82)$$

The values of j can be

$$j = |\ell - s|, |\ell - s| + 1, \dots, (\ell + s) - 1, (\ell + s) \quad (3.83)$$

3.5.2 Relativistic kinetic energy

The classical Hamilton $H = T + V$ is only accurate for small momenta, and using accurate measurement techniques there will always be a discrepancy between results obtained using $T + V$ and results using the relativistic Hamiltonian: $H = \sqrt{m^2 c^4 + c^2 \|\mathbf{p} - e\mathbf{A}(\mathbf{x}, t)\|^2} + e\Psi(\mathbf{x}, t)$. By Taylor approximating this Hamiltonian for low velocities we get

$$H \approx T + V + O\left(\frac{v^2}{c^2}\right) \quad (3.84)$$

Hence the relativistic correction we make would be of order $v^2 c^{-2}$. We find that the ratio vc^{-1} is

$$\frac{v}{c} = \frac{\alpha}{n}, \quad \frac{e^2}{4\pi\epsilon_0} \frac{1}{\hbar c} \quad (3.85)$$

where α is the **fine structure constant**, and $\alpha \approx \frac{1}{137}$. Therefore, the relativistic correction is of order α^2 , as are all other fine structure corrections.

3.5.3 Darwin Term

For electrons with $\ell = 0$ there is a non-zero probability to be at (and near) the site of the nucleus, where the energy density is so high that we must take into account the effects described by the Dirac equation. The Dirac equation predicts that there will be a *Zitterbewegung*, due to electron-positron pairs created from the vacuum (vacuum fluctuations). The interaction of the electron with this EP-pair causes the *Zitterbewegung*. This correction is proportional to $|\psi(r=0)|^2$, and therefore only affects s-electrons

3.5.4 Lamb shift

3.6 HELIUM

The Hamiltonian that describes a Helium atom is

$$H = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) + \frac{e^2}{4\pi\epsilon_0} \left(-\frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}} \right) \quad (3.86)$$

where Ze is the nuclear charge and $r_{12} = \frac{1}{\|\mathbf{r}_1 - \mathbf{r}_2\|}$. We begin by neglecting the electron-electron interaction; the Hamiltonian uncouples and the total wave function is the product of hydrogenic electron wave functions. This means the total energy would be

$$E^{(0)} = E_1 + E_2 \approx -109\text{eV}, \quad \text{ignoring repulsion} \quad (3.87)$$

where I have used that $Z = 2$ for Helium and hence $E = -4\hbar c R_\infty$. Now we treat the electron-electron repulsion as a perturbation:

$$E^{(1)} = \langle V_{12} \rangle = \frac{e^2}{4\pi\epsilon_0} \langle (1s)^2 | r_{12}^{-1} | (1s)^2 \rangle = 34\text{eV} \quad (3.88)$$

This implies that it is not appropriate to treat the electron-electron repulsion as a perturbation. However adding this value gives us an energy in the right region:

$$E^{(0)} + E^{(1)} = -75\text{eV} \quad (3.89)$$

whereas the experimentally measured value is -78.95eV .

3.6.1 Excited states

*we can't tell the difference
between $(1s)(n\ell)$ and
 $(n\ell)(1s)$.*

For excited states we follow the same procedure, however using different wave functions. We assume that only one electron is excited, hence we look at the configuration $(1s)(n\ell)$. However, electrons are indistinguishable, therefore there is an **exchange degeneracy**, which implies we should use degenerate perturbation theory. Let us look at the wave function:

$$|\psi\rangle = a|(1s)(n\ell)\rangle + b|(n\ell)(1s)\rangle \quad (3.90)$$

Using which we can construct the perturbation matrix:

$$\begin{pmatrix} J & K \\ K & J \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \Delta E \begin{pmatrix} a \\ b \end{pmatrix} \quad (3.91)$$

where

$$J \equiv \frac{e^2}{4\pi\epsilon_0} \langle (1s)(n\ell) | \frac{1}{r_{12}} | (1s)(n\ell) \rangle \quad (3.92)$$

$$K \equiv \frac{e^2}{4\pi\epsilon_0} \langle (1s)(n\ell) | \frac{1}{r_{12}} | (n\ell)(1s) \rangle \quad (3.93)$$

This gives us a correction of

$$\Delta E = J \pm K \quad (3.94)$$

with the following eigenstates:

$$|\psi_{\text{space}}^S\rangle = \frac{1}{\sqrt{2}} (|(1s)(n\ell)\rangle + |(n\ell)(1s)\rangle) \quad (3.95)$$

$$|\psi_{\text{space}}^A\rangle = \frac{1}{\sqrt{2}} (|(1s)(n\ell)\rangle - |(n\ell)(1s)\rangle) \quad (3.96)$$

The symmetric wave function has energy $E^{(0)} + \Delta E$, whereas the antisymmetric wave function has $E^{(0)} - \Delta E$.

3.6.2 Spin states

Fermions have antisymmetric wave functions under particle-label interchange. This implies that the total wave function (space times spin) must be of the form

$$|\Psi\rangle = \begin{cases} |\psi_{\text{space}}^S \psi_{\text{spin}}^A\rangle \\ |\psi_{\text{space}}^A \psi_{\text{spin}}^S\rangle \end{cases} \quad (3.97)$$

Note that there only one antisymmetric spin configuration, forming a singlet:

$$|\psi_{\text{spin}}^A\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \quad (3.98)$$

which has $S = M_S = 0$. The symmetric spin configuration forms a triplet

$$|\psi_{\text{spin}}^S\rangle = \begin{cases} |\uparrow\uparrow\rangle \\ \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |\downarrow\downarrow\rangle \end{cases} \quad (3.99)$$

all with $S = 1$ and $M_S = 1, 0, -1$ respectively. Using spectroscopic notation, the terms (^{2S+1}L) of the ground state would be 1S and 3S for the singlet and triplet respectively.

3.7 ALKALIS

The alkalis are the first group of the periodic system, hence they have one electron in the outermost shell, and all inner shells are filled. For example

$$\text{Lithium : } (1s)^2 2s \quad (3.100)$$

$$\text{Sodium : } (1s)^2 (2s)^2 (2p)^6 3s \quad (3.101)$$

$$\text{Potassium : } (1s)^2 (2s)^2 (2p)^6 (3s)^2 (3p)^6 4s \quad (3.102)$$

$$\vdots \quad (3.103)$$

3.7.1 The quantum defect

The $3s$, $3p$ and $3d$ configurations all have the same energy for hydrogen, however this is not the case for atoms with more than one electron – the electron-electron repulsion breaks this degeneracy. This is in part due to the fact that the inner electrons “screen” the nucleus – the outermost electrons see a nucleus with a charge $< Ze$. This shows why the $3s$ has a different energy than the $3p$ configuration: $3s$ has a higher probability to be closer to the nucleus, where it experiences a larger effective charge. Despite this complication, Bohr’s formula works well, as long as we introduce a new parametre:

$$E(n, \ell) = -hc \frac{R_\infty}{(n - \delta_\ell)} \quad (3.104)$$

where δ_ℓ is the *quantum defect*. For example sodium has the following quantum defects:

$$\delta_s = 1.35, \quad \delta_p = 0.86, \quad \delta_d = 0.01, \quad \delta_\ell \approx 0 \text{ for } \ell > 2 \quad (3.105)$$

The decrease of δ_ℓ with increasing ℓ is because the higher ℓ the smaller the probability is of the electron being inside the electron cloud of the closed shells.

3.7.2 Central-field approximation

The result from the previous example is purely empirical, and can’t tell us anything more than what we have measured experimentally. In order to, for example, find the wave functions for the

outermost electrons in alkalis, we must approximate the Schrödinger equation. Commonly one uses the central field approximation, which uses the spherical symmetry of closed subshells. This repulsion between subshell-electron and the outermost electron is shoved into a central potential, $S(r)$. Note that we assume that all electron-electron interactions are spherically symmetric, which is why this approximation is only realistic for the alkalis. We define the central-field potential

$$V_{\text{CF}}(r) = -\frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r} + S(r) \quad (3.106)$$

and hence the central-field Hamiltonian is

$$H_{\text{CF}} = \sum_{i=1}^N \left(-\frac{\hbar^2}{2m} \nabla_i^2 + V_{\text{CF}}(r_i) \right) \quad (3.107)$$

this is a separable Hamiltonian, hence

$$\psi_{\text{atom}} = \prod_{i=1}^N \psi_i, \quad E = \sum_{i=1}^N E_i \quad (3.108)$$

where

$$\left(-\frac{\hbar^2}{2m} \nabla_i^2 + V_{\text{CF}}(r_i) \right) \psi_i = E_i \psi_i \quad (3.109)$$

The central potential can be thought of as a spherically symmetric potential, where the effective nuclear charge eZ_{eff} is a function of the radius, due to shielding. The radial dependence of this effective nuclear charge can be estimated by treating the subshell-electrons as hydrogenic electrons, and then evaluating the total electric field due to the closed subshell together with the nucleus.

3.7.3 Self-consistent solutions

Once we go through the calculations of the central field potential, we realise that the potential changing will cause our wave-functions to change, which in turn causes the potential to change etc. What one can do, numerically, is repeat this process over and over again, until the corrections one makes between each iteration are negligible. Once this is done one has made a self-consistent solution.

3.7.4 Fine structure in the alkalis

The fine structure of the alkalis can be approximated quite well using the Landé formula:

$$\Delta E_{\text{FS}} = \frac{Z_i^2 Z_o}{(n + \delta_\ell)^3 \ell(\ell + 1)} \alpha^2 \hbar c R_\infty \quad (3.110)$$

Where Z_i and Z_o are the inner and outer atomic numbers respectively. For neutral alkali atoms $Z_o \approx 1$ and generally $Z_i \sim Z$, because no shielding occurs on the inside of the atom. This formula can be justified by looking at how we would evaluate the fine structure splitting using the central-field approximation. We would need to evaluate the value of

$$\left\langle \frac{Z_{\text{eff}}(r)}{r^3} \right\rangle = \left\langle \frac{1}{er} \frac{\partial V_{\text{CF}}(r)}{\partial r} \right\rangle \quad (3.111)$$

which indeed results in the Landé formula.

3.8 LS- AND jj-COUPLING SCHEME

In reality the Hamiltonian is not spherically symmetric and in fact the complete Hamiltonian cannot be written with just the central-field term, therefore we should write the full Hamiltonian as

$$H = H_{\text{CF}} + H_{\text{RE}} \quad (3.112)$$

where H_{RE} is the residual electrostatic interaction between the electrons. This interaction would be of the form

$$H_{\text{RE}} = \sum_i \sum_{j>i} \left(\frac{e^2}{4\pi\epsilon_0 r_{ij}} - S(r_i) \right) \quad (3.113)$$

The residual electrostatic interaction causes the i and j -th electrons' angular momentum to precess (ℓ_i and ℓ_j interact with each other), therefore these are no longer good quantum numbers. However, the residual interaction results in an internal force/torque, which cannot change the total angular momentum $\mathbf{L} = \ell_1 + \ell_2$. H_{RE} does not affect the spin, hence $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$ is still a good quantum number. Thus $|S M_S L M_L\rangle$ are the good quantum states.

Good quantum numbers correspond to classical conserved quantities.

If $E_{\text{SO}} \ll E_{\text{RE}}$ we can treat the spin-orbit interaction as a perturbation, and use the states $|S M_S L M_L\rangle$ as the unperturbed states, however, the perturbation affects the z -components of S and M_L , which no longer are constant, which is why it is beneficial to use the coupled basis $|S L J M_J\rangle$ – this is what is referred to as the **LS-coupling scheme**. In this scheme we speak of **terms**, which is just what we call the states where we have specified L and S . For example $(3p)(4p)$ in silicon. Here

$$\ell_1 = \ell_2 = 1 \rightarrow L \in \{0, 1, 2\} \quad (3.114)$$

$$s_1 = s_2 = \frac{1}{2} \rightarrow S \in \{0, 1\} \quad (3.115)$$

We denote each of these states with the **term**, ^{2S+1}L , where we replace $L = 1$ with S , $L = 2$ with P , $L = 3$ with D etc. Thus we get six terms that correspond to the configuration

$$(3p)(4p) \rightsquigarrow ^1S, ^1P, ^1D, ^3S, ^3P, ^3D \quad (3.116)$$

$2S + 1$ is the multiplicity of the term; for example 1S is a singlet state, whereas 3P is a triplet state.

An important point to make here, which is relevant only for *equivalent electrons*, is that the Pauli exclusion principle requires the total wave function of electrons to be anti-symmetric under label exchange. The spatial wave functions that give an odd numbered angular momentum L are anti-symmetric under label exchange, and hence must be multiplied with a symmetric spin state (for example the triplets that have $S = 1$). This holds generally and it can be shown that

$$L + S \equiv 0 \pmod{2}, \quad \text{for equivalent electrons} \quad (3.117)$$

For example one would expect that the ground state configuration of helium $(1s)^2$ has the following terms:

$$(1s)^2 \rightsquigarrow ^1S, ^3S \quad (3.118)$$

because $\ell_1 = \ell_2 = 0$ and $s_1 = s_2 = \frac{1}{2}$. However the greyed out state has an odd numbered $L + S$ and hence it can quickly be seen that this state is not allowed by Pauli's exclusion principle.

3.8.1 Spin-orbit interaction in the LS-coupling scheme

The spin-orbit interaction for each electron is summed up for the total interaction:

$$H_{\text{SO}} = \beta_1 \mathbf{s}_1 \cdot \boldsymbol{\ell}_1 + \beta_2 \mathbf{s}_2 \cdot \boldsymbol{\ell}_2 \quad (3.119)$$

however, \mathbf{s}_i and $\boldsymbol{\ell}_i$ precess about \mathbf{S} and \mathbf{L} respectively. \mathbf{S} and \mathbf{L} in turn precess about $\mathbf{J} = \mathbf{L} + \mathbf{S}$, however in the LS-coupling scheme this precession is slow, which is why we can use L and S . Using the projection theorem we can write the spin-orbit interaction in terms of the total spin and angular momenta:

$$H_{\text{SO}} = \beta_{\text{SO}} \mathbf{L} \cdot \mathbf{S} \quad (3.120)$$

From here we can express the spin-orbit energy as a function of J , L and S :

$$E_{\text{SO}} = \frac{\beta_{\text{SO}}}{2} (J(J+1) - L(L+1) - S(S+1)) \quad (3.121)$$

where $J \in \{|L-S|, |L-S|+1, \dots, L+S-1, L+S\}$.

The difference between two adjacent J values' spin orbit energies is given by the **interval rule**:

$$\Delta E_{\text{SO}} = E_J - E_{J-1} = \beta_{\text{SO}} J \quad (3.122)$$

The spin-orbit interaction lifts the J degeneracy, which means we should now additionally specify the J value, this turns a term into a **level**, $^{2S+1}L_J$. For example the 3P term of ground state silicon splits into three level:

$$^3P \rightsquigarrow ^3P_0, ^3P_1, ^3P_2 \quad (3.123)$$

However, we know that, due to the residual electrostatic interaction and due to spin-orbit coupling, that the degeneracy in L and J is lifted. The question now is: how do these levels' energies order themselves, with H_{RE} and H_{SO} taken into account? The answer is given by Hund's rules. However, this rule is only accurate for *equivalent electrons* in the *ground state*, and hence cannot be used for excited states. These are **Hund's rules**:

1. The lowest energy state has largest S consistent with Pauli's exclusion principle
2. If there are several of these the lowest energy term has the largest value of L .
3. For a given term in an atom with outermost sub-shell half-filled or less, the level with the lowest value of J will have the lowest energy level. Otherwise it is the level with the greatest value of J that has the lowest energy.

On the other side of the scale would be where $E_{\text{SO}} \gg E_{\text{RE}}$, in which case we need to use the spin-orbit interactions eigenstates as the unperturbed eigenstates. This is generally the case for heavy atoms. Here \mathbf{s}_i and $\boldsymbol{\ell}_i$ couple to \mathbf{j}_i . Here it is common to denote the levels as $(j_1, j_2)_J$, for example an sp configuration splits into four levels:

$$(sp) \rightsquigarrow \left(\frac{1}{2}, \frac{1}{2}\right)_0, \left(\frac{1}{2}, \frac{1}{2}\right)_1, \left(\frac{1}{2}, \frac{3}{2}\right)_1, \left(\frac{1}{2}, \frac{3}{2}\right)_2 \quad (3.124)$$

In conclusion:

$$\text{LS-coupling scheme: } E_{\text{SO}} \ll E_{\text{RE}}$$

$$\text{jj-coupling scheme: } E_{\text{SO}} \gg E_{\text{RE}}$$

The projection theorem states that, because $\boldsymbol{\ell}_i$ precesses quickly about \mathbf{L} , we need only consider the (expectation value of the) projection of $\boldsymbol{\ell}_i$ onto \mathbf{L} : $\langle \boldsymbol{\ell}_i \cdot \mathbf{L} \rangle$

3.8.2 Selection rules in the LS-coupling scheme:

Selection Rule	Exception
$\Delta J = 0, \pm 1$	$J = 0 \leftrightarrow J' = 0$
$\Delta M_J = 0, \pm 1$	$M_J = 0 \leftrightarrow M'_J = 0$ if $\Delta J = 0$
Parity changes	
$\Delta \ell = \pm 1$	One electron jump
$\Delta L = 0, \pm 1$	$L = 0 \leftrightarrow L' = 0$
$\Delta S = 0$	

Table 3.1: Selection rules for electric dipole transitions.

3.8.3 Zeeman effect in the LS-coupling scheme

In this scheme we can treat the Zeeman effect for monovalent atoms. The magnetic moment of an atom is

neglecting the nuclear magnetic moment

$$\boldsymbol{\mu} = -\mu_B \mathbf{L} - g_s \mu_B \mathbf{S} \quad (3.125)$$

where the spin g-factor $g_s = 2$ according to Dirac, however QED increases that value by about one per mille. The *anomalous* Zeeman effect is the effect where the second term $g_s \mu_B \mathbf{S}$ contributes. If we can treat the Zeeman effect as a perturbation to the spin-orbit coupling, which in the LS-coupling scheme is a perturbation to the residual electrostatic scheme, then we can use $|L S J M_J\rangle$ as our unperturbed states. The energy relations are:

$$E_{Ze} \ll E_{SO} \ll E_{RE} \quad (3.126)$$

The interaction energy between a magnetic moment and a magnetic field is $H = -\boldsymbol{\mu} \cdot \mathbf{B}$. Once again using the projection theorem and assuming that $\mathbf{B} \parallel \mathbf{z}$ we get

$$H_{Ze} = \frac{\langle \mathbf{L} \cdot \mathbf{J} \rangle + g_s \langle \mathbf{S} \cdot \mathbf{J} \rangle}{J(J+1)} \mu_B B J_z \rightsquigarrow E_{Ze} = g_J \mu_B B J_z \quad (3.127)$$

where:

assuming $g_s = 2$, exactly.

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \quad (3.128)$$

for singlets $g_J = 1$. Additionally singlets have $S = 0$, therefore these experience the normal Zeeman effect.

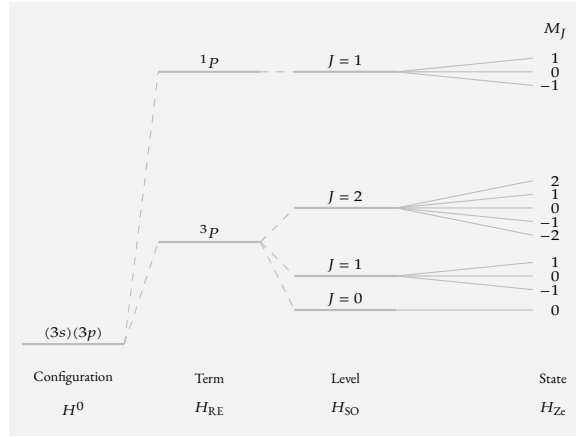


Figure 3.2: The hierarchy of atomic structure for the (3s)(3p) configuration of an alkaline earth metal atom.

3.9 HYPERFINE STRUCTURE AND ISOTOPE SHIFT

The nucleus has a magnetic moment:

$$\mu_I = g_I \mu_N \mathbf{I} \quad (3.129)$$

where $\mu_N \approx \frac{\mu_B}{1836}$ is the nuclear magneton. The hyperfine structure is given by the interaction between the nuclear and electronic magnetic moments:

$$H_{HFS} = -\mu_I \cdot \mathbf{B} = A \mathbf{I} \cdot \mathbf{J} \quad (3.130)$$

where A is the hyperfine constant. Similar to what we did for the spin-orbit interaction we define the total angular momentum $\mathbf{F} = \mathbf{I} + \mathbf{J}$. The hyperfine interaction causes \mathbf{I} and \mathbf{J} to precess about \mathbf{F} , which means that their magnitudes are constant, however, their z -components are not- just like how it was for the spin-orbit interaction. Thus we go from the uncoupled basis, $|I M_I J M_J\rangle$ to the coupled basis $|I J F M_F\rangle$, this is sometimes referred to as the IJ -coupling scheme. This means that

$$E_{HFS} = \frac{A}{2} (F(F+1) - I(I+1) - J(J+1)) \quad (3.131)$$

Here we also have an interval rule, exactly like with the spin-orbit interaction:

$$\Delta E_{HFS} = E_F - E_{F-1} = AF \quad (3.132)$$

3.9.1 Isotope shift

There are two contributions towards isotope splitting. The first we have already encountered in the Bohr model: the mass shift (because the nucleus does not have infinite mass). The second contribution is due to the nonzero volume of the nucleus.

Mass shift

As with the Bohr model we can perform the effective-mass substitution: $m_e \rightsquigarrow \frac{m_e M_N}{m_e + M_N}$, which gives us a true wavenumber

$$\tilde{\nu} = \tilde{\nu}_\infty \left(\frac{M_N}{m_e + M_N} \right) \quad (3.133)$$

where M_N is the nuclear mass. However $\tilde{\nu}_\infty$ is not measurable. But we *can* measure is the difference wavenumbers between two isotopes, say with nuclear masses $A'M_p$ and $A''M_p$ respectively. This gives us

$$\Delta\tilde{\nu}_{\text{Mass}} = \tilde{\nu}_{A'} - \tilde{\nu}_{A''} \approx \frac{m_e}{M_p} \frac{A' - A''}{A'A''} \tilde{\nu}_\infty \quad (3.134)$$

For light atoms this correction is relevant. For example it is greater than the fine structure in Hydrogen.

We see from Equation 3.133 that the heavier isotope always has larger wavenumbers.

Volume shift

The nucleus has a nonzero volume, which will cause a shift for s-electrons. The nuclear charge distribution, ρ_N , will interact with the electric potential due to the electron, ϕ_e :

$$E_{\text{Vol}} = \int d\mathbf{r} \rho_N \phi_e = \frac{Ze^2}{6\epsilon_0} |\psi(0)|^2 \langle r_N^2 \rangle \quad (3.135)$$

The "liquid drop model" gives a formula for the radius of the nucleus:

$$r_N \approx 1.2A^{\frac{1}{3}} \text{ fm} \quad (3.136)$$

where AM_p is the nuclear mass. This gives us an expression for the isotope shift:

$$\Delta\tilde{\nu}_{\text{Vol}} = \frac{\Delta E_{\text{Vol}}}{hc} \approx \frac{\langle r_N^2 \rangle}{a_0^2} \frac{\Delta A}{A} \frac{Z^2}{(n + \delta_\ell)^3} R_\infty \quad (3.137)$$

Using the mass shift and the volume shift, we can, using photonic excitations draw conclusions about the structure of nuclei.

3.9.2 Zeeman effect and hyperfine structure

We will treat the Zeeman effect for weak, intermediate and strong magnetic fields. The strength of the magnetic field is compared to the hyperfine structure constant, $\frac{A}{\mu_B}$.

The magnetic moment of an atom is

$$\boldsymbol{\mu}_{\text{atom}} = -g_J\mu_B\mathbf{J} + g_I\mu_N\mathbf{I} \approx -g_J\mu_B\mathbf{J} \quad (3.138)$$

We can neglect the nuclear magnetic moment and still be accurate within about one per mille, as $\mu_B \approx 1836\mu_N$. This gives us the interaction energy for the Zeeman effect:

$$H_{Ze} = g_J\mu_B\mathbf{J} \cdot \mathbf{B} \quad (3.139)$$

Weak-field Zeeman effect, $\mu_B B < A$.

In this limit the hyperfine interaction is stronger than the Zeeman effect, therefore we should use the IJ -coupling scheme. This is because \mathbf{I} and \mathbf{J} precess rapidly about \mathbf{F} , whereas \mathbf{F} precesses slowly about \mathbf{B} . Hence we should use the projection theorem once again:

$$H = g_J\mu_B \frac{\langle \mathbf{J} \cdot \mathbf{F} \rangle}{F(F+1)} \mathbf{F} \cdot \mathbf{B} = g_F\mu_B B F_z \quad (3.140)$$

where

$$g_F = \frac{F(F+1) + J(J+1) - I(I+1)}{2F(F+1)} g_J \quad (3.141)$$

This gives us

$$E = g_F \mu_B B M_F \quad (3.142)$$

Intermediate-field Zeeman effect, $\mu_B B \approx A$

In this case one treats $H_{Ze} + H_{HFS}$ as a perturbation. In this limit one can still use the IJ -coupling scheme, as the unperturbed basis. For example the ground state of hydrogen would have the following perturbation matrix:

$$H' = \begin{pmatrix} \frac{A}{4} & -\mu_B B \\ -\mu_B B & -\frac{3A}{4} \end{pmatrix} \quad (3.143)$$

which gives a splitting of

$$E' = -\frac{A}{4} \pm \sqrt{\frac{A^2}{4} + \mu_B^2 B^2} \quad (3.144)$$

which is a hyperbola in B . This is the splitting for the two $M_F = 0$ states. The $M_F = \pm 1$ states have a linear splitting.

Strong-field Zeeman effect, $\mu_B B > A$

In this regime the IJ -coupling scheme is no longer applicable, because F no longer is a good quantum number (\mathbf{I} and \mathbf{J} precess quickly and independently about \mathbf{B}). Instead it can be shown that

$$E = g_J \mu_B B M_J + A M_I M_J \quad (3.145)$$

The quick precession of \mathbf{I} and \mathbf{J} about \mathbf{B} cause the x - and y -components to average out.

3.9.3 Measurement of the Zeeman effect

Optical techniques are, generally, not suitable for the measurement of the Zeeman effect, because the Zeeman effect is of the same order of magnitude as the Doppler broadening. Instead one can use the atomic-beam technique, which in principle is two Stern-Gerlach apparatus in succession. The first apparatus gathers the for example $M_J = \pm \frac{1}{2}$ states. Then the a radio frequency apparatus is used to induce transitions. The second apparatus is the either

1. used to gather the states again (by having a magnetic field gradient that points in the same direction as the first apparatus' gradient). This would be a "flop-in" arrangement.
2. used to deflect the states (by having a magnetic field gradient that points in the opposite direction to the first apparatus' gradient). This would be a "flop-out" arrangement.

Generally it is true that the flop-in arrangement measures no flux, unless the radio-frequency waves have the correct frequency (corresponding to the hyperfine energy splitting), whereas the flop-out arrangement measures flux only when the radio-frequency waves have the correct frequency.

3.IO CAESIUM ATOMIC CLOCKS

The hyperfine splitting of Caesium splits the level into $F = 3$ and $F = 4$ states. Atomic clocks use the transition that has the smallest sensitivity to magnetic fields, which is the transition

$$|F = 4, M_F = 0\rangle \leftrightarrow |F = 3, M_F = 0\rangle \quad (3.146)$$

For weak magnetic fields, usually in the μT region (atomic clocks usually use magnetic fields with about this strength) the $M_F = 0$ states feel no Zeeman splitting, however $M_F \neq 0$ states experience a linear Zeeman effect of the order $\Delta f_B = 6.998 M_F \cdot B \frac{\text{MHz}}{\text{T}}$.

Atomic clocks have two regions (A and B) where the magnetic field is non-uniform (say it varies along the z -axis). This magnetic field gradient gives rise to a force on the atoms:

$$F_{\text{atom}} = -\frac{\partial W}{\partial z} = -\frac{\partial W}{\partial B} \frac{\partial B}{\partial z} = -\mu_{\text{eff}} \frac{\partial B}{\partial z} \quad (3.147)$$

note that μ_{eff} is as function of M_F . At non-zero magnetic fields the states $|F = 4, M_F = 0\rangle$ and $|F = 3, M_F = 0\rangle$ will experience an opposite force (the $|F = 3, M_F = 0\rangle$ state is a high-field seeker, as its energy is reduced for increasing magnetic fields). The polariser field (A -field) is used to pick out either $|F = 4, M_F = 0\rangle$ or $|F = 3, M_F = 0\rangle$.

The atoms are then sent through a C -region with a homogenous magnetic field, where the atoms are allowed to evolve naturally. We will get to the C -region momentarily.

Then the atoms fly through the B -field, also known as the analyser, where we do the same thing as in the A -region, but with an opposite magnetic field. This means that if the evolution truly is free in the C -region, then we will not see any atoms hit the detector.

If we now use a Ramsey cavity within the C -region, which will drive the transition $|4, 0\rangle \leftrightarrow |3, 0\rangle$, using two $\pi/2$ -pulses. Beginning in the north pole of the Bloch Sphere the first $\pi/2$ -pulse takes us to the equatorial-plane, where we evolve naturally for some time.

If we are in the rotating frame, then this free evolution has phase $e^{i\delta t}$. The second $\pi/2$ -pulse will take us to the south pole if and only if $e^{i\delta T} = 0$ which is true if $T = \frac{2\pi}{\delta}$ or $\delta = 0$. The former is extremely unlikely, therefore we can use this method to determine whether the radio synthesiser has zero detuning— whether or not $\omega = \omega_0$.

Scanning through frequencies in the frequency synthesiser will give us a Ramsey resonance structure, which allows us to very precisely pin-point the value of ω_0 .

It is very important that the two $\pi/2$ -pulses are temporally-coherent.

3.II DIATOMIC MOLECULES

3.II.1 Energy contributions

Diatomic molecules have more degrees of freedom than atoms, however, the additional degrees of freedom do not affect the gross structure, but rather fine tune the energy scales. That is to say the gross structure is still due to electronic energies. The additional degrees of freedom are two

rotational degrees of freedom and two vibrational degree of freedom ($T + V$). Here are the energy scales:

1. Electronic energies: $E_e \sim \frac{\hbar^2}{ma^2} \sim \text{eV}$
2. Vibrational energies: $E_v \sim \sqrt{\frac{m_e}{M}} E_e \sim 0.1 \text{eV}$
3. Rotational energies: $E_r \sim \frac{m_e}{M} E_e \sim \text{meV}$

where M is the molecular mass.

The full problem that we need to solve is:

$$[T_N + T_e + V]\Psi = E\Psi \quad (3.148)$$

where we use reduced mass of the nuclei $\mu = \frac{M_A M_B}{M_A + M_B}$ and define $\mathbf{R} = \mathbf{R}_A - \mathbf{R}_B$ to be the distance between nuclei:

$$T_N = -\frac{\hbar^2}{2\mu} \nabla_{\mathbf{R}}^2 \quad (3.149)$$

$$T_e = -\frac{\hbar^2}{2m_e} \sum_i \nabla_{\mathbf{r}_i}^2 \quad (3.150)$$

$$V = \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z_A Z_B}{R} + \sum_i \left(-\frac{Z_A}{\|\mathbf{r}_i - \mathbf{R}_A\|} \right. \right. \quad (3.151)$$

$$\left. \left. -\frac{Z_B}{\|\mathbf{r}_i - \mathbf{R}_B\|} + \sum_{j \neq i} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} \right) \right) \quad (3.152)$$

because $\mu \gg m_e$ we can use the **Born-Oppenheimer approximation**:

1. "Freeze" the nuclei; solve the electronic wave-functions, ϕ_q :

$$(T_e + V(\mathbf{R}, \mathbf{r}_i)) \phi_q(\mathbf{R}, \mathbf{r}_i) = E_q(\mathbf{R}) \phi_q(\mathbf{R}, \mathbf{r}_i) \quad (3.153)$$

2. We treat \mathbf{R} as a variation-parametre and minimize $E_q(\mathbf{R})$. $\phi_q(\mathbf{R}, \mathbf{r}_i)$ form a complete set for each internuclear distance. Solve the Schrödinger equation for the nuclear part of the wave function with $E_q(\mathbf{R})$ as the potential.

The potential energy surfaces $E_q(R)$ have minima, which is the most convenient internuclear distance. We taylor expand the potential energy surfaces, assuming $R \approx R_0$:

This should be an accurate approximation for normal temperatures.

$$E_q(R) \approx E_q(R_0) + \frac{1}{2}k(R - R_0)^2 \quad (3.154)$$

Treating this quantum mechanically will give us harmonic oscillator energy solutions:

$$E_v = \hbar\omega_0 \left(v + \frac{1}{2} \right), \quad \omega_0 = \sqrt{\frac{k}{\mu}} \quad (3.155)$$

The rotational part is of the form $\frac{\hbar^2}{2\mu} \frac{J(J+1)}{R_0^2}$ where we have assumed that $R \approx R_0$, thus the total energy is

$$E_{q,v,J} = E_q(R_0) + \hbar\omega_0 \left(v + \frac{1}{2} \right) + \frac{\hbar^2}{2\mu R_0^2} J(J+1) \quad (3.156)$$

for low energies. That higher excitations the nuclear potential can no longer be treated as a harmonic oscillator.

This leads to the **Franck-Condon principle**, which is a weak "selection rule". The electronic transitions occur on very fast time scales compared to the nuclear motion, thus the nuclear coordinates barely change during the transition. Thus we *favour* transitions for which $\Delta R = 0$. Or rather we favour transitions whose radial overlap is largest:

$$f_{v,v'}^{q,q'} = \int \left(\psi_{v'}^{q'} \right)^* \psi_v^q dR \quad (3.157)$$

is known as the Franck-Condon factor, and the transition strength is proportional to this factor.

3.11.2 Diatomic electric dipole transitions

The harmonic oscillator approximation gives us selections rules:

$$\Delta v = \pm 1 \quad (3.158)$$

and the quantised angular momentum gives us an additional selection rule:

$$\Delta J = \pm 1 \quad (3.159)$$

This defines two sets of transition, during *absorption*:

1. *R* transitions, for which $\Delta J = 1$.
2. *P* transitions, for which $\Delta J = -1$.

these sets make up the *R*-branch and *P*-branch respectively.

3.12 DOPPLER FREE LASER SPECTROSCOPY

Doppler broadening, which will be discussed shortly causes a line-width, which often is not negligible, for fine structure, but especially for hyperfine structure. Therefore numerous methods have been developed to avoid the effect of Doppler broadening.

3.12.1 Doppler broadening

The velocity distribution of atoms in an ideal gas is given by the Maxwell distribution. This gives us the fraction of atoms in the range from v to $v + dv$:

$$f(v)dv = \frac{1}{u\sqrt{\pi}} \exp\left(-\frac{v^2}{u^2}\right) dv, \quad u \equiv \sqrt{\frac{2k_B T}{M}} \quad (3.160)$$

where u is the most probable speed for atoms at temperature T . The Doppler effect says that a reference frame moving with velocity v with respect to a stationary reference frame, which sends out light with frequency ω , will measure a shifted frequency, ω' , equal to

$$\omega' = \omega - \mathbf{k} \cdot \mathbf{v} \quad (3.161)$$

where \mathbf{k} is the wave vector of the transmitted light. This gives us a Gaussian line shape function:

$$g_D(\omega) = \frac{c}{u\omega_0\sqrt{\pi}} \exp\left(-\frac{c^2}{u^2} \left(\frac{\omega - \omega_0}{\omega_0}\right)^2\right) \quad (3.162)$$

where ω_0 is the resonant frequency (atomic transition frequency). This Gaussian has a FWHM of

$$\Delta\omega_D = \left(2\sqrt{\ln 2} \frac{u}{c}\right) \omega_0 = 2\sqrt{\ln 2} \frac{u}{\lambda} \quad (3.163)$$

3.12.2 The crossed-beam method

A simple way to reduce Doppler broadening, is by minimising the component of the velocity of the gas that is parallel to the laser beam used for measurements. This can be done by sending the laser beam at a right angle to an atomic beam. The atomic beam is not collinear, however, if we send the atomic beam through a small slit at a large distance from the oven we can reduce the collimation angle significantly. For small collimation angles $\sin \alpha \approx \alpha$, hence

$$\Delta f \sim \alpha \Delta f_D \quad (3.164)$$

where Δf_D is the normal Doppler broadening of the gas at the same temperature. Suppose we would like the Doppler broadening of Sodium vapour at 1000K to be comparable to the natural line width, in that case

$$\alpha \Delta f_D \approx \Delta f_N \rightarrow \alpha \approx 4 \cdot 10^{-3} \text{rad} \quad (3.165)$$

where we have used that $v_{\text{beam}} \approx 1000 \text{m s}^{-1}$. This corresponds to a right angle where the ratio between opposite and adjacent is about $\frac{1}{250}$, which is doable in a laboratory environment.

3.12.3 Saturated Absorption Spectroscopy

Suppose we have a gas of atoms, with some Doppler-broadening ω_D and natural transition frequency ω_0 . We now begin by sending an intense $I \approx I_{\text{sat}}$ laser beam with frequency ω , through the atoms (call this the *pump beam*). This will promote about half the population in the velocity class $v = (\omega - \omega_0)/k$ to the excited level. This promotion will have a homogeneous power-broadening line width.

Now we send another laser beam through the atoms with low intensity $I \ll I_{\text{sat}}$ in the opposite direction (call this the *probe beam*). This will promote a few atoms in the velocity class $v = -(\omega - \omega_0)/k$ assuming that $\omega \neq \omega_0$ – some of the laser beam will be absorbed. However, assuming now that $\omega \approx \omega_0$ there will be about as much excitation as there is de-excitation, due to the pump beam. Thus there will be general shape of increasing absorption as $\omega \rightarrow \omega_0$, however, when $\omega \approx \omega_0$ there will be a sudden drop in absorption.

Cross-over resonances in saturation spectroscopy

The discussion above is only accurate for two-level systems. In reality there are more than one transition that can occur. Naturally each of which will occur at different frequencies. Suppose we have a three-level system with transition frequencies ω_1 and ω_2 . Our spectroscopy will naturally

have absorption drops at these two frequencies, however, there will be an addition drop, namely there we

$$\frac{\omega_1 + \omega_2}{2} = \omega \quad (3.166)$$

This is exactly the point when the ω_1 excitations from the pump beam affect the ω_2 excitations of the probe beam and vice versa. In general there will be additional drops at

$$\frac{\omega_i + \omega_j}{2} = \omega, \quad i \neq j \quad (3.167)$$

There are $\frac{N(N-1)}{2}$ additional drops, if there are N true transitions. However note that, depending on what the nature of the transition is, some of these additional drops can lie within each other. For hyperfine structure the interval rule holds, therefore for small enough N you will be able to see all drops, however some may be in the form of double-drops.

3.12.4 Two-photon spectroscopy

By sending two counter-propagating photons with $\omega = \frac{\omega_0}{2}$ through the atomic gas the Doppler broadening will be cancelled completely:

$$\omega \left(1 + \frac{v}{c}\right) + \omega \left(1 - \frac{v}{c}\right) = 2\omega = \omega_0 \quad (3.168)$$

The benefit of this vs saturation spectroscopy is that all atoms can absorb the two photons, not just those at rest in the laboratory frame. Therefore we can expect stronger signal and hence better measurements.

3.13 LASER COOLING AND TRAPPING

3.13.1 Scattering Force

Photons bear momentum, which is transferred to atoms during interactions. The force is equal to how much momentum is transferred per unit time, which equals rate at which the light delivers energy divided by the speed of light:

$$F_{\text{rad}} = \frac{IA}{c} \sim 10^{-9}\text{N}, \quad \text{for } IA \sim 1\text{W} \quad (3.169)$$

The area, A can be taken as $\sigma(\omega)$ per atom. At its peak ($\omega \approx \omega_0$) the radiative cross-section is much greater than the size of atoms, which implies that F_{rad} can also spike around resonance. We needn't think about the effect of the atom having to give spontaneously emitted photons a part of their momentum, because this effect will average out.

The magnitude of this *scattering* force equals the rate at which absorb photons impart momentum to the atom:

$$F_{\text{scatt}} = \hbar k \frac{\Gamma}{2} \frac{I/I_{\text{sat}}}{1 + I/I_{\text{sat}} + 4\delta^2/\Gamma^2} \quad (3.170)$$

which gives a maximum acceleration of

$$a_{\max} = \frac{\hbar k \Gamma}{2M} = \frac{v_r \Gamma}{2} \quad (3.171)$$

where we have defined that recoil velocity $v_r = \frac{\hbar k}{M} \cdot a_{\max} \sim 10^5 \text{ m s}^{-2}$ for sodium, which is much greater than the gravitational acceleration. We can define the stopping distance (ignoring any Doppler-effects):

$$L_0 = \frac{v_0^2}{a_{\max}} \quad (3.172)$$

for $v_0 = 1000 \text{ m s}^{-1}$ the stopping distance is $\sim 1 \text{ m}$.

3.13.2 *Slowing an atomic beam*

In practice one cannot neglect the Doppler-shift, which changes as a function of z (assuming that the atoms propagate in the z -direction). One method uses the Zeeman effect to vary the transition-frequency as a function of z , by using a varying magnetic field. Using the assumption that the deceleration is constant we get that

$$v(z) = v_0 \left(1 - \frac{z}{L_0} \right)^{\frac{1}{2}} \quad (3.173)$$

We would like for the Doppler corrected frequency $\omega - kv(z)$ to match with the transition frequency ω_0 , this we can do by using the Zeeman effect, which perturbs the transition frequency $\omega_0 + \frac{\mu_B B}{\hbar}$, hence our condition is

$$\omega_0 + \frac{\mu_B B(z)}{\hbar} = \omega + kv(z) \rightsquigarrow B(z) = B_0 \left(1 - \frac{z}{L_0} \right)^{\frac{1}{2}} + B_{\text{bias}} \quad (3.174)$$

where

$$B_0 = \frac{\hbar k v_0}{\mu_B} \quad (3.175)$$

3.13.3 *Chirp Cooling*

Instead of adjusting the transition frequency as a function of z one can adjust the laser frequency as a function of time – we change the laser frequency such that the laser signal is always Doppler-corrected.

3.13.4 *Optical Molasses Technique*

Suppose we have an atom moving at some velocity v . If we now send light with frequency ω from both the left and the right then the atom will experience a force

$$F_{\text{molasses}} = F_{\text{scatt}}(\omega - \omega_0 - kv) - F_{\text{scatt}}(\omega - \omega_0 + kv) \quad (3.176)$$

This is, by definition an odd function of v . A first order Taylor approximation gives us

$$F_{\text{molasses}} \approx -2kv \frac{\partial F_{\text{scatt}}}{\partial \omega} = -\alpha v \quad (3.177)$$

where

$$\alpha = 2kv \frac{\partial F_{\text{scatt}}}{\partial \omega} = 4\hbar k^2 \frac{I}{I_{\text{sat}}} \frac{-2\delta/\Gamma}{(1 + (2\delta/\Gamma)^2)^2} \quad (3.178)$$

Not the same α as used previously... sorry

which is positive only for $\delta < 0$ which is equivalent to the condition

$$\omega < \omega_0 \quad (3.179)$$

thus this technique only works for light that is red-shifted with respect to the atomic transition frequency. We want $\alpha > 0$ so that the molasses force points in the opposite direction than v . Note that we have once again neglected I/I_{sat} in the denominator, because we assume that $I/I_{\text{sat}} \ll 1$.

This gives us an exponential decay of energy:

$$\frac{dE}{dt} = -\frac{2\beta}{M}E = -\frac{E}{\tau_{\text{damp}}} \quad (3.180)$$

However this does not continue forever, the Doppler-cooling techniques do have limits, which will be discussed shortly.

3.13.5 Doppler cooling limit

The momentum kicks during spontaneous emission as well as the randomness of the absorption are source of uncertainty in both the position and the momentum of the atoms. We find that Doppler cooling has temperature limit, T_D :

$$T_D = \frac{\hbar\Gamma}{2k_B} \quad (3.181)$$

the *Doppler cooling limit*. For sodium this temperature is $T_D = 240\mu\text{K}$. Though in reality these techniques work better than expected.

3.13.6 Magneto-optical trap

The optical molasses technique could hold atoms in place in outer space, however down on earth we need to overcome gravity. This is done using the magneto-optical trap, which combines the optical molasses technique (optical) with an anti-Helmholtz magnetic field. However, here the laser beams have circular polarisation, and it has a frequency that is slightly less than the atomic resonance frequency $J = 0 \rightsquigarrow J = 1$ in fact slightly less than the lowest of these transitions (there are three).

In the middle of the coils the magnetic fields cancel and there is $B = 0$. In the vicinity of this point however, there is a uniform field gradient, which perturbs the atomic energies, causing the three $J = 1$ states' energies to vary linearly with their displacement from the middle. The Zeeman effect

causes an energy shift $\omega_0 \pm \beta z$. The Doppler broadening makes the atom see an effective frequency of $\omega \pm kv$.

This σ^+ -light comes from the bottom and σ^- comes from the top. Thus for $z < 0$ the $|0\ 0\rangle \leftrightarrow |1\ 1\rangle$ will occur, whereas for $z > 0$ the $|0\ 0\rangle \leftrightarrow |1\ -1\rangle$ transition will occur. Due to the uniform magnetic-field-gradient the Zeeman shift will vary with z , hence the MOT force is:

$$F_{\text{MOT}} = F_{\text{scatt}}^{\sigma^+}(\omega - kv - (\omega + \beta z)) - F_{\text{scatt}}^{\sigma^-}(\omega + kv - (\omega - \beta z)) \quad (3.182)$$

$$\approx -2kv \frac{\partial F}{\partial \omega} + 2\beta z \frac{\partial F}{\partial \omega_0} \quad (3.183)$$

The $\omega \pm kv$ term is due to the Doppler Effect whereas the $\omega \pm \beta z$ is due to the Zeeman Effect

Hence we have a dampened harmonic oscillator, which will hold atoms in place, also in the presence of gravity (gravity will add a constant to the differential equation, just shift the equilibrium position slightly)

3.A APPENDIX

3.A.I Summary of atomic units

Bohr radius:

$$a_0 = \frac{\hbar^2}{m_e} \frac{4\pi\epsilon_0}{e^2} = 5.29 \cdot 10^{-11} \text{m} \quad (3.184)$$

Hydrogen ground-state energy:

$$hcR_\infty = \frac{m_e}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 = 13.6 \text{eV} \quad (3.185)$$

Electron-volt to Joule conversion

$$1 \text{eV} = 1.602 \cdot 10^{-19} \text{J} \quad (3.186)$$

Zeeman frequency shift, due to magnetic field:

$$\frac{\Omega_L}{2\pi B} = \frac{e}{4\pi m_e} = 14 \text{GHz T}^{-1} \quad (3.187)$$

Bohr magneton:

$$\mu_B = \frac{e\hbar}{2m_e} = 9.27 \cdot 10^{-24} \text{J T}^{-1} \quad (3.188)$$

4.1 SPECIFIC HEAT OF SOLIDS

As far back as 1819 it was discovered that for many solids the heat capacity per atom is

$$C = 3k_B, \quad k_B = 1.38 \cdot 10^{-23} \frac{\text{J}}{\text{K}} \quad (4.1)$$

This is true for many monoatomic solids, however, there are a few exceptions, such as diamond. This was later explained quite accurately, by Stefan Boltzmann (in 1896), using a classical model, treating each atom as a harmonic oscillator. However, this law could not sufficiently describe the low temperature behaviour of solids.

4.1.1 *Einstein Model of solids*

Einstein's idea was similar to that of Boltzmann, however, he treated the harmonic oscillators quantum mechanically. The n -th eigenenergy of the harmonic oscillator is

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad (4.2)$$

where ω is the eigenfrequency of the harmonic oscillator, sometimes referred to as the Einstein frequency. Using this we can define the partition function:

$$Z = \sum_n e^{-\beta\hbar\omega(n+\frac{1}{2})} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} = \frac{1}{2 \sinh(\beta\hbar\omega/2)} \quad (4.3)$$

From here we can define most thermodynamic quantities, such as, for example the average energy

$$\langle E \rangle = \hbar\omega \left(n_B(\beta\hbar\omega) + \frac{1}{2} \right) \quad (4.4)$$

where the Bose-Einstein distribution is given by

$$n_B(x) = \frac{1}{e^x - 1} \quad (4.5)$$

This is easy to generalise to three dimensions ($n + \frac{1}{2} \rightsquigarrow (n_x + \frac{1}{2}) + (n_y + \frac{1}{2}) + (n_z + \frac{1}{2})$), using the rules for the exponential function:

$$Z_{3D} = (Z_{1D})^3 \quad (4.6)$$

The derivative of the average energy with respect to temperature gives us the heat capacity:

$$C = 3k_B (\beta\hbar\omega)^2 \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} \quad (4.7)$$

This means the heat capacity is effectively constant for high temperatures, compared to the Einstein temperature $T_{\text{Einstein}} \equiv \frac{\hbar\omega}{k_B}$, and decays exponentially for low temperatures:

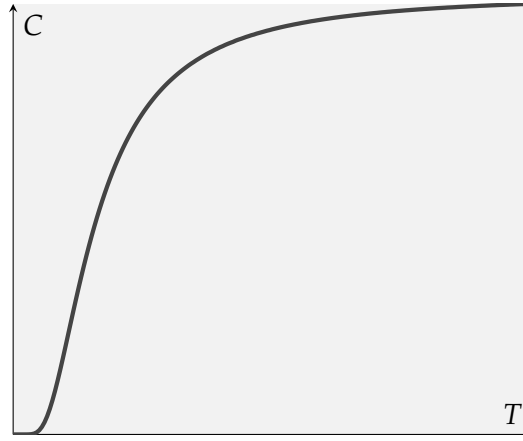


Figure 4.1: Heat capacity as a function of T .

Note that we recover $C = 3k_B$ for high temperatures, and that the vibrational degree of freedom "freezes out" at low temperatures. For most materials the Einstein temperature is low compared to room temperature, however, for diamond $T_{\text{Einstein}} \approx 1320\text{K}$, which is why $C < 3k_B$ per atom for Diamond. This is due to the strong atomic bonds between carbon atoms.

4.1.2 Debye Model of solids

Einstein's result did quite well, however, only qualitatively in the low temperature region. It was shown experimentally that there would be a T^3 dependence for low temperatures, which is not at all what Einstein's calculations predicted.

Debye realised that the vibrations inside solids are waves, and that waves should be quantised similarly to how Planck quantised light in 1900. However, there are a few differences between light and sound. Naturally the speed of sound is far lower than the speed of light, but additionally light only has two polarisations for any given \mathbf{k} , whereas sound has three vibrational modes (two transversal and one longitudinal).

More often than not we will look at waves using *periodic boundary conditions*. For a sample length of L this restricts the possible values of k to

$$k = \frac{2\pi n}{L}, \quad n \in \mathbb{Z} \quad (4.8)$$

If we are required to sum over all k values, we can instead integrate as follows:

$$\sum_k \rightsquigarrow \frac{L}{2\pi} \int_{-\infty}^{\infty} dk \quad (4.9)$$

Generally though, in three dimensions this looks quite similar:

$$\mathbf{k} = \frac{2\pi}{L} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \quad n_1, n_2, n_3 \in \mathbb{Z} \quad (4.10)$$

and the summation can also be replaced by a volumetric integral in k -space:

$$\sum_{\mathbf{k}} \rightsquigarrow \left(\frac{L}{2\pi}\right)^3 \int d\mathbf{k} \quad (4.11)$$

It's often beneficial to use spherical coordinates to carry out this integral.

Debye *assumed* a linear dispersion relation:

$$\omega(\mathbf{k}) = v\|\mathbf{k}\| \quad (4.12)$$

This assumption turns out to be accurate for long wavelengths, however not for shorter wavelengths, as we will see.

Now we would like to sum over all values of \mathbf{k} :

$$\langle E \rangle = 3 \sum_{\mathbf{k}} \hbar \omega(\mathbf{k}) \left(n_B(\beta \hbar \omega(\mathbf{k})) + \frac{1}{2} \right) \quad (4.13)$$

$$= \frac{3L^3}{(2\pi)^3} \int d\mathbf{k} \hbar \omega(\mathbf{k}) \left(n_B(\beta \hbar \omega(\mathbf{k})) + \frac{1}{2} \right) \quad (4.14)$$

Using spherical symmetry we convert this into a spherical integral, which we additionally can simplify, by using $k = \frac{\omega}{v}$, to

$$\langle E \rangle = \int_0^\infty d\omega g(\omega) (\hbar \omega) \left(n_B(\beta \hbar \omega) + \frac{1}{2} \right) \quad (4.15)$$

where we have taken everything from the integral that is not $\varepsilon(\omega)$ and combined it into the **density of states**:

$$g(\omega) = \frac{9N\omega^2}{\omega_D^3}, \quad \omega_D = 6\pi^2 n v^3 \quad (4.16)$$

where ω_D is known as the Debye frequency. The term with $\frac{1}{2}$ integrates up to a temperature independent constant, therefore it does not contribute to the heat capacity. We get

Which unfortunately is infinite.

$$\langle E \rangle = \frac{9N\hbar}{\omega_D^3 (\beta \hbar)^4} \int_0^\infty dx \frac{x^3}{e^x - 1} + C \quad (4.17)$$

where C is the temperature independent constant. The integral is just a constant, so we get

$$\langle E \rangle = \frac{9N\hbar}{\omega_D^3 (\beta \hbar)^4} \frac{\pi^4}{15} + C \quad (4.18)$$

This gives us an expression for the heat capacity:

$$C = \frac{\partial \langle E \rangle}{\partial T} = N k_B \frac{T^3}{T_D} \frac{12\pi^4}{5}, \quad T_D \equiv \frac{\hbar \omega_D}{k_B} \quad (4.19)$$

This gives us the T^3 dependence, as expected. Unfortunately this expression is proportional to T^3 for all temperatures, however, this is because it is not reasonable to assume that all sound modes are activated (it is certainly not the case that we have an infinite number of sound modes, all with increasing energy). Debye fixed this problem by introducing a cutoff frequency:

$$3N = \int_0^{\omega_{\text{cutoff}}} d\omega g(\omega) \quad (4.20)$$

Using this expression for the cutoff frequency, we can set an upper limit for the frequency in our integral for the average energy. At high temperatures the Bose-Einstein distribution becomes $\approx \frac{k_B T}{\hbar \omega}$, so we get

$$\langle E \rangle = k_B T \int_0^{\omega_{\text{cutoff}}} d\omega g(\omega) = 3Nk_B T \quad (4.21)$$

which is exactly Dulong-Petit. Evaluating the cutoff frequency:

$$3N = \int_0^{\omega_{\text{cutoff}}} d\omega g(\omega) = 9N \int_0^{\omega_{\text{cutoff}}} \frac{\omega^2}{\omega_D^3} = 3N \frac{\omega_{\text{cutoff}}^3}{\omega_D^3} \quad (4.22)$$

hence the cutoff frequency is exactly the Debye frequency. Interestingly enough $k_D = \frac{\omega_D}{v} = (6\pi^2 n)^{\frac{1}{3}}$ is of the order of inverse interatomic distance.

Debye's theory has a few shortcomings:

1. The cutoff frequency is more of a hack than good physics.
2. The linear dispersion is not accurate, especially for large value of k .
3. Debye theory is not accurate at intermediate temperatures.
4. This model cannot explain the linear temperature dependent term that appears in metals in the low temperature limit.

4.2 VIBRATIONS OF A ONE-DIMENSIONAL MONATOMIC CHAIN

Consider a harmonic monatomic chain with spring constant κ and lattice constant a . Denote the position of the n -th atom by x_n and its equilibrium position as x_n^{eq} . Using this we can define the n -th atom's displacement

$$\delta x_n = x_n - x_n^{\text{eq}} \quad (4.23)$$

The total potential is a sum over the individual harmonic oscillators:

$$V_{\text{tot}} = \sum_i V(x_{i+1} - x_i) = \sum_i \frac{\kappa}{2} (\delta x_{i+1} - \delta x_i)^2 \quad (4.24)$$

There are only two terms that depend on δx_n , so luckily the force $F = -\nabla V = m\delta\ddot{x}_n$ is evaluated quite easily:

$$m\delta\ddot{x}_n = \kappa (\delta x_{n+1} + \delta x_{n-1} - 2\delta x_n) \quad (4.25)$$

We will now use the Ansatz, that the oscillations will be of the form:

$$\delta x_n = A e^{i(\omega t - k x_n^{\text{eq}})} = A e^{i(\omega t - k n a)} \quad (4.26)$$

where I have used that the spacing between atoms is a and assumed that $x_0^{\text{eq}} = 0$. This gives us

$$-m\omega^2 A e^{i(\omega t - k n a)} = \kappa A e^{i(\omega t - k n a)} (e^{-ika} + e^{ika} - 2) \quad (4.27)$$

hence

$$\omega^2 = 2\omega_0^2 (1 - \cos(ka)) = 4\omega_0^2 \sin^2\left(\frac{ka}{2}\right), \quad \omega_0^2 = \frac{\kappa}{m} \quad (4.28)$$

Thus we obtain

$$\omega = 2\omega_0 \left| \sin \left(\frac{ka}{2} \right) \right| \quad (4.29)$$

which looks like this

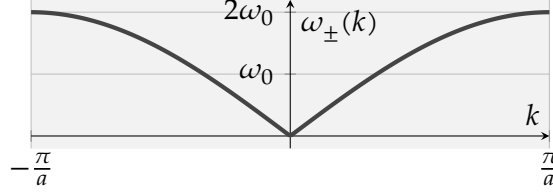


Figure 4.2: The first Brillouin zone for a monatomic harmonic chain.

Notice that $\omega(k)$ is periodic in k , so there is no difference between $\omega(k)$ and $\omega(k + 2\pi/a)$. This is actually far more general than it appears: there is no physical difference between modes with k and $k + 2\pi/a$, hence we can look at the Brillouin zone alone, which is defined as the values of $|k|$ that are not equivalent to a smaller value of $|k|$. This is related to residue classes, as

$$k \equiv k + G \pmod{G}, \quad G = \frac{2\pi n}{a}, \quad n \in \mathbb{Z} \quad (4.30)$$

hence we just need to look at all different $[k]$ – the Brillouin zone.

This leads us to a very general and important result:

If a lattice is periodic in real space with periodicity a , then the reciprocal lattice is periodic, with periodicity $2\pi/a$.

For example this implies that if the unit cell has width a then the Brillouin zone will have width $2\pi/a$.

Given a direct lattice with lattice points denoted by x_n we can find the reciprocal lattice points, G_m by

$$e^{iG_m x_n} = 1 \quad (4.31)$$

4.2.1 Properties of the one-dimensional chain

The velocity of wave packets is defined by the *group velocity*, which follows from Hamiltonian mechanics:

$$v_{\text{group}} \equiv \frac{d\omega}{dk} \quad (4.32)$$

additionally we can define the *phase velocity* as

$$v_{\text{phase}} \equiv \frac{\omega}{k} \quad (4.33)$$

which is the velocity at which individual maxima and minima. Using our dispersion relation we see that long wavelength waves (small values of k) have an approximately constant group velocity

(these are sound waves). At higher values of k the group velocity decreases, and it reaches zero at the Brillouin zone boundary.

The total number of normal modes is the size of the Brillouin zone ($2\pi/a$) divided by the spacing between neighbouring k values ($2\pi/(Na)$), hence there are N normal modes, which is precisely what Debye *assumed* previously.

4.2.2 Quantum Modes: Phonons

Let us introduce the quantum correspondence principle:

If a classical harmonic system has a normal mode with frequency ω then the corresponding quantum system will have eigenstates with energy

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad (4.34)$$

This fits very well with what found in QM with the quantum harmonic oscillator, however, this also holds for coupled harmonic oscillators. This leads us to define the **phonon**: a quantum of vibration. There is no constraint to how many phonons can occupy the same energy simultaneously, hence phonons are bosons. Therefore the occupancy of any mode is described with the Bose-Einstein distribution:

$$n_B(\beta\hbar\omega) = \frac{1}{e^{\beta\hbar\omega} + 1} \quad (4.35)$$

and hence the expectation value of the energy is

$$\langle E_k \rangle = \hbar\omega(k) \left(n_B(\beta\hbar\omega) + \frac{1}{2} \right) \quad (4.36)$$

The total energy is given by the sum of these individual terms, which can be evaluated by an integral instead, as discussed previously. We get

$$U = \frac{Na}{2\pi} \int_{-\pi/a}^{\pi/a} dk \hbar\omega(k) \left(n_B(\beta\hbar\omega) + \frac{1}{2} \right) \quad (4.37)$$

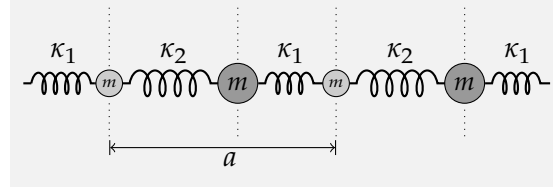
using a change in variables this integral is written as an integral with respect to ω , where we need to include the density of states over frequency, as we did previously. However now we can see that

$$g(\omega) = \frac{Na}{\pi} \left| \frac{dk}{d\omega} \right| \quad (4.38)$$

where the extra factor of two is because each value of ω corresponds to two values of k (one positive and one negative).

4.3 VIBRATIONS OF A ONE-DIMENSIONAL DIATOMIC CHAIN

Let us now instead consider a one-dimensional diatomic chain. They will have equal masses, say m , and two different spring constants, say κ_1 and κ_2 . As in the following figure



We could from here define a *basis*, which is where each atom is with respect to the unit cell. For example the basis for the diatomic chain in the figure above would be

light gray atom at the position of the lattice point

dark gray atom $\frac{17}{30}a$ to the right of the lattice point.

4.3.I Normal modes of the diatomic chain

The calculation for a diatomic chain is very similar to that of the monoatomic chain, however, now we have *two* displacements we need to keep track of. Let us denote the displacement of the light gray atoms by δx_n and the dark gray atoms by δy_n . We would then get

$$m\delta\ddot{x}_n = \kappa_2 (\delta y_n - \delta x_n) + \kappa_1 (\delta y_{n-1} - \delta x_n) \quad (4.39)$$

$$m\delta\ddot{y}_n = \kappa_1 (\delta x_{n+1} - \delta y_n) + \kappa_2 (\delta x_n - \delta y_n) \quad (4.40)$$

Once again assuming planar waves we get that

$$\delta x_n = \delta x_{n+1} e^{ika} \quad (4.41)$$

$$\delta y_n = \delta y_{n-1} e^{-ika} \quad (4.42)$$

and also that $\delta\ddot{x}_n = \omega^2 \delta x_n$, thus

$$m\omega^2 \begin{pmatrix} \delta x_n \\ \delta y_n \end{pmatrix} = \begin{pmatrix} \kappa_1 + \kappa_2 & -\kappa_2 - \kappa_1 e^{ika} \\ -\kappa_2 - \kappa_1 e^{-ika} & \kappa_1 + \kappa_2 \end{pmatrix} \begin{pmatrix} \delta x_n \\ \delta y_n \end{pmatrix} \quad (4.43)$$

The eigenvalue equation gives us

$$\omega_{\pm} = \sqrt{\frac{\kappa_1 + \kappa_2}{m} \pm \frac{1}{m} \sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1\kappa_2 \cos(ka)}} \quad (4.44)$$

Which looks like

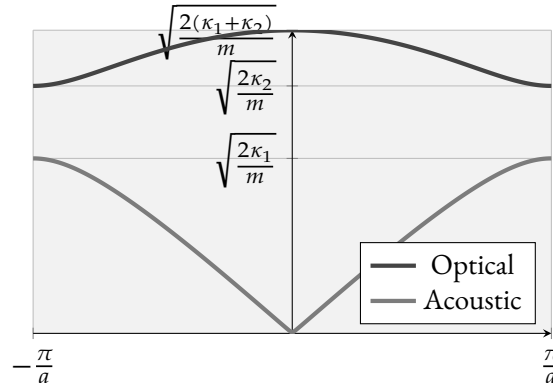


Figure 4.3: Dispersion relation for the diatomic chain in the reduced scheme.

For $k \rightarrow 0$ on the acoustic mode we have the sound waves, for which there is a linear dispersion. We can define the speed of sound in this limit:

$$v_{\text{sound}} = \frac{d\omega_-}{dk} = \sqrt{\frac{a^2 \kappa_1 \kappa_2}{2m(\kappa_1 + \kappa_2)}} \quad (4.45)$$

4.4 ELECTRONS IN METALS

4.4.1 Drude theory

The first model of electrons discussed in the course is the Drude theory, which makes the following assumptions:

1. Electrons have a scattering time τ , which corresponds to the average time between collisions. The probability of scattering in the time interval dt is dt/τ .
2. Once a scattering event occurs we assume the momentum returns to $\mathbf{p} = \mathbf{0}$. This is not true for individual electrons, but is the case for the expectation value of \mathbf{p} .
3. In between scattering events electrons respond to external electric and magnetic fields as per the Lorentz force.

Using this we can derive a differential equation that relates the momentum to the applied electric and magnetic fields – the Drude equation:

$$\frac{d\mathbf{p}}{dt} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) - \frac{\mathbf{p}}{\tau} \quad (4.46)$$

At equilibrium we get that

$$E = \left(\frac{\mathbf{j} \times \mathbf{B}}{ne} + \frac{m}{ne^2\tau} \mathbf{j} \right) \quad (4.47)$$

where n is the electron density and we have used that $\mathbf{j} = -nev$. Using this we can define the **resistivity tensor**, ρ_{ij} , whose "diagonal" entries are the Drude resistivity:

$$\rho_{ii} = \frac{m}{ne^2\tau} \quad (4.48)$$

For $\mathbf{B} \parallel \mathbf{z}$ we get

$$\rho_{xy} = -\rho_{yx} = \frac{B}{ne} \quad (4.49)$$

the Hall resistivity. The Hall coefficient R_H is defined as

$$R_H = \frac{\rho_{yx}}{|\mathbf{B}|} = -\frac{1}{ne} \quad (4.50)$$

where the final equality is the result from Drude theory. We can define the **conductivity tensor**, σ_{ij} as the inverse of ρ_{ij} , hence

$$j_\alpha = \sigma_{\alpha\beta} E_\beta, \quad \text{RISO} \quad (4.51)$$

RISO means repeated indices are summed over.

4.4.2 Thermal transport

The Drude theory of electrons, together the kinetic theory of gas can help explain the Wiedemann-Franz law. κ is the thermal conductivity, which is used in the diffusion equation. The ratio $\frac{\kappa}{T\sigma} \equiv L$ is known as the Lorenz number, and the two theories together predict that

$$L = \frac{\kappa}{T\sigma} = \frac{3}{2} \left(\frac{k_B}{e} \right)^2 \approx 1.11 \cdot 10^{-8} \text{W}\Omega\text{K}^{-2} \quad (4.52)$$

Wiedemann and Franz found that the ratio $\frac{\kappa}{T\sigma}$ is approximately constant for most metals.

which is very close (up to about a factor two) to the experimentally observed values for various metals.

However, this theory is wrong, but we get approximately the right results due to two mistakes that luckily cancel. We have vastly underestimated the velocity of electrons in metals, and vastly overestimated the specific heat of electrons in metals (both of which were used to calculate the heat conductance). In conclusion: we cannot use classical gas theory to describe electrons in a metal.

This can be seen when looking at the Seebeck and Peltier effects. The Peltier effect states that electric currents also cause heat currents:

$$\mathbf{j}^q = \Pi \mathbf{j} \quad (4.53)$$

where Π is the Peltier coefficient. In kinetic theory of gasses the thermal current is

$$\mathbf{j}^q = \frac{1}{3} c_V T n \mathbf{v} \quad (4.54)$$

where c_V is the specific heat. This leads to a Seebeck coefficient of

$$S = \frac{\Pi}{T} = \frac{-k_B}{2e} = -0.43 \cdot 10^{-4} \text{V K}^{-1} \quad (4.55)$$

which is off by about a factor for 100 for most metals, and in fact some metals have a Seebeck coefficient that has the opposite sign.

4.4.3 Sommerfeld theory

Sommerfeld theory incorporates Fermi-Dirac statistics, into the description of electrons in metals. Given a system of non-interacting electrons with chemical potential, μ , the probability of an eigenstate with energy E being occupied is given by the Fermi-Dirac distribution:

$$n_F(\beta(E - \mu)) = \frac{1}{e^{\beta(E - \mu)} + 1} \quad (4.56)$$

Note the similarity to the Bose-Einstein distribution. The difference being the opposite sign in the denominator.

Now consider a box with volume $V = L^3$ with periodic boundary conditions. The eigenstates will be planar waves, which quantised wave vectors:

$$\varepsilon(\mathbf{k}) = \frac{\hbar^2 \|\mathbf{k}\|^2}{2m}, \quad \mathbf{k} = \frac{2\pi}{L} \begin{pmatrix} n \\ m \\ \ell \end{pmatrix}, \quad n, m, \ell \in \mathbb{Z} \quad (4.57)$$

It doesn't really matter whether we use periodic boundary conditions or fixed boundary conditions, as the main contributor to the physics of the system is the bulk, of which there is much more than there are borders.

At $T = 0$ there Fermi-Dirac distribution is a reversed Heavyside function, which flips at the Fermi energy:

$$n_F(\beta(E - \mu)) \xrightarrow{T \rightarrow 0} \Theta(E_F - \varepsilon(\mathbf{k})) \quad (4.58)$$

Where the Fermi energy is defined as $E_F = \mu(T = 0)$. The Heavyside function leads to a natural definition of the Fermi wavevector:

$$E_F - \varepsilon(\mathbf{k}) = 0 \rightsquigarrow E_F = \frac{\hbar^2 k_F^2}{2m} \quad (4.59)$$

At $T = 0$ we can integrate n_F with respect to \mathbf{k} to get the total number of electrons, and use the spherical symmetry to get

$$N = \frac{V}{4\pi^3} \left(\frac{4}{3} \pi k_F^3 \right) \rightsquigarrow E_F = \frac{\hbar^2 (3\pi^2 n)^{\frac{2}{3}}}{2m} \quad (4.60)$$

Using the relation between ε and k we can change any integral over k into an integral over ε , for example

$$E_{\text{total}} = V \int_0^\infty d\varepsilon \varepsilon g(\varepsilon) n_F(\beta(\varepsilon - \mu)) \quad (4.61)$$

where $g(\varepsilon)$ is the density of states per unit volume, over energy. Unfortunately it is very difficult to calculate this integral for non-zero temperatures, due to the Fermi-Dirac distribution. The **Sommerfeld expansion** uses the fact that the integral is easily calculated at $T = 0$ and Taylor expands the integral from there, giving

$$E(T) = E(T = 0) + \frac{\tilde{\gamma}}{2} (Vg(E_F)(k_B T)) (k_B T) + \dots \quad (4.62)$$

where $\tilde{\gamma} = \frac{\pi}{3}$. Which gives us a linear heat capacity

$$C = \tilde{\gamma} \left(\frac{3Nk_B}{2} \right) \left(\frac{T}{T_F} \right) \quad (4.63)$$

which is the missing linear term for the heat capacity of metals – the linear terms is hence because of the electrons in the metal, that do not behave classically.

Unfortunately this model too has its shortcomings, which is why we shall venture even further. Here are some things that are not explained by Sommerfeld theory:

1. Why aren't electrons scattered by the positively charged nuclei?
2. Why do core electrons not contribute to the Fermi energy and velocity, and why don't insulators have any free electrons?
3. We still do not understand why the Hall coefficient sometimes has the wrong sign.
4. Many optical features, such as high harmonic absorption are not explained by the Sommerfeld theory.
5. Sometimes measurements of the specific heat of electrons is still off by a factor 10. Additionally we sometimes measure the wrong electron mass
6. We cannot explain magnetism
7. Electron interaction – we ignored electron-electrons interactions, and there is no reason this should be accurate.

4.5 CHEMICAL BONDS

4.5.1 Covalent Bond

Suppose we have two hydrogen atoms that are close enough to each other for them to interact. We will use the *Born-Oppenheimer approximation*, which states that we can view the nuclei as stationary, and solve for the electronic wave functions separately. Once this is done one would use the electronic energy surfaces as a function of the internuclear distancing to solve the nuclear wave functions, however, that is not what we are doing now.

Let us consider a single electron between the nuclei. This will have the following Hamiltonian

$$H = T + V_1 + V_2, \quad V_i = -\frac{e^2}{4\pi\epsilon_0\|\mathbf{r} - \mathbf{R}_i\|} \quad (4.64)$$

Using the variational principle we will take a trial wave-function:

$$|\psi\rangle = \phi_1|1\rangle + \phi_2|2\rangle \quad (4.65)$$

where

$$(T + V_1)|1\rangle = \epsilon_0|1\rangle \quad (4.66)$$

$$(T + V_2)|2\rangle = \epsilon_0|2\rangle \quad (4.67)$$

assume $\langle i|j\rangle = \delta_{ij}$ and define

$$\langle 1|V_2|1\rangle = \langle 2|V_1|2\rangle = V_{\text{cross}} \quad (4.68)$$

where the first equality is due to the indistinguishability of the two nuclei. Additionally let us define the *hopping* term:

$$\langle 1|V_1|2\rangle = t \quad (4.69)$$

let us assume $t > 0$. Now this gives us a Schrödinger equation of the form:

$$\begin{pmatrix} \epsilon_0 + V_{\text{cross}} & t \\ t & \epsilon_0 + V_{\text{cross}} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = E \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \quad (4.70)$$

which has the following eigenvalues:

$$E_{\pm} = \epsilon_0 + V_{\text{cross}} \pm t \quad (4.71)$$

with eigenvectors

$$|\psi_{\text{bonding}}\rangle = \frac{1}{\sqrt{2}} (|1\rangle - |2\rangle) \quad (4.72)$$

$$|\psi_{\text{antibonding}}\rangle = \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle) \quad (4.73)$$

This says that there exists a state $|\phi_{\text{bonding}}\rangle$ where it is more convenient for the hydrogen nuclei to be bonded together, hence forming H^+ , as this has lower energy than the energy of the two individual hydrogen nuclei with the electron on one of them.

This is the basic principle behind covalent bonding.

4.5.2 *Van der Waals, Fluctuating Dipole Forces, or Molecular Bonding*

Atoms have a dipole moment, whose expectation is zero. However, the dipole moment can fluctuate momentarily, which will induce a dipole-force between atoms. This is what is known as the Van der Waals force. It is weak compared to ionic and covalent bonds, however the force can act over a further distance, because electrons don't need to hop.

4.5.3 *Metallic Bonding*

It is sometimes difficult to distinguish between covalent and metallic bonds – both involve the sharing of electrons through delocalisation. However for the metallic bonds the delocalisation is on the scale of the entire crystal, whereas covalent delocalisation only happens between two atoms (or a few). The delocalised electrons act as free particles, which conduct electricity. The delocalisation does not favour any direction particularly, therefore slightly moving atoms will not affect the bonds – metals are often ductile and malleable.

4.5.4 *Hydrogen Bonding*

The exceptionally neutral electronegativity and small size of hydrogen means that hydrogen can take on both a positive and a negative partial charge. This special case is why we treat hydrogen bonding separately to other atomic bonds. The partial charges of hydrogen and the donator/acceptor cause a dipole-moment, which attracts neighbouring molecules. In the example of water the oxygen atoms take on a negative partial charge, which attracts neighbouring positively (and partially) charged hydrogen atoms.

In summary:

1. **Ionic bonds:** Electron is transferred from one atom to another. Formed by atoms with very different electronegativity. Hard, brittle, high melting temperature, insulator, hydrophilic.
2. **Covalent bonds:** Electron is shared between neighbouring atoms. Usually atoms in the groups *III* to *V*. A great example is silicon and diamonds. Very hard, high melting temperature, insulator or semiconductor.
3. **Metallic bonds:** Electrons are delocalised throughout the lattice. Left and middle of the periodic table. Ductile, malleable, hardened by impurities, lower melting temperature, good electrical and thermal conductors.
4. **Van der Waals bonding:** No transfer of electrons, only due to the oscillating dipole moment of atoms. Noble gas solids, non-polar solids. Soft, weak, low melting temperature, insulator.
5. **Hydrogen bonding:** Involves partial charged hydrogen atoms. Important in organic materials. Holds together ice. Weak bond (stronger than Van der Waals).

4.6 TIGHT BINDING CHAIN

4.6.1 1D chain

Let us denote the wavefunction of electron at the n -th electron as $|n\rangle$. As we saw previously the interactions between these atoms will cause the electron orbits to hybridise. Hence the hybridised wave function will be some linear combination of atomic orbitals:

$$|\Psi\rangle = \sum_n \phi_n |n\rangle, \quad \langle n|m\rangle = \delta_{nm} \quad (4.74)$$

where the orthonormality is an *assumption*, the orbitals at each site needn't be orthogonal, and in fact aren't orthogonal for small interatomic distances. The time-independent Schrödinger equation in this Hilbert space is

$$\sum_m H_{nm} \phi_m = E \phi_n, \quad \langle \phi_n | H | \phi_m \rangle \quad (4.75)$$

Let us say that the electrons have an on-site energy of ε_0 , and a hopping energy of t :

$$H_{nm} = \varepsilon_0 \delta_{nm} + t (\delta_{n-1,m} + \delta_{n+1,m}) \quad (4.76)$$

Let us now assume that our wave function is of the form:

$$\phi_n = \frac{e^{-ikna}}{\sqrt{N}} \quad (4.77)$$

plugging this into our time-independent Schrödinger equation gives us

$$E = \varepsilon + 2t \cos(ka) \quad (4.78)$$

note that t is normally negative, hence it is convenient for the electrons to dissociate. This means its convenient to write

$$E = \varepsilon - 2|t| \cos(ka) \quad (4.79)$$

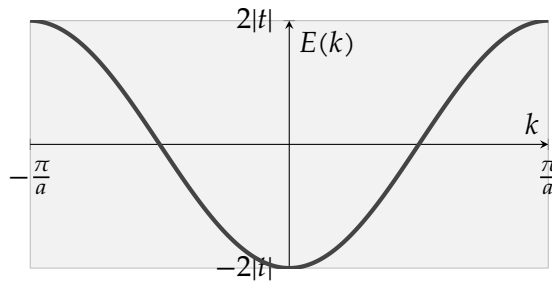


Figure 4.4: The first Brillouin zone for a monatomic tight-binding chain.

Note that in the long wavelength limit this is parabolic:

$$E \approx E_0 + \hbar^2 k^2 / 2m^* \quad (4.80)$$

where $E_0 = \varepsilon - 2|t|$. This means that our linear combination of atomic orbitals behave as a quasiparticle, whose mass is

$$\frac{\hbar^2 k^2}{2m^*} = \hbar^2 k^2 / 2ta^2 \quad \rightsquigarrow \quad m^* = \frac{\hbar^2}{2ta^2} \quad (4.81)$$

4.6.2 *Filling bands*

In the ground state $T = 0$ we fill the states from the bottom upwards. However, Pauli's exclusion principle says that we can fill each state with two electron (one with up and one with down spin). Thus for monovalent atoms we would only fill half the band:

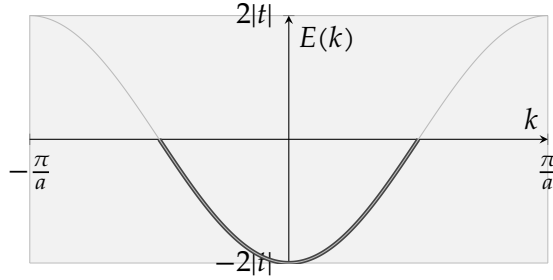


Figure 4.5: The first Brillouin zone for a monatomic, monovalent tight-binding chain.

Applying a constant electric field would apply a force $F = -eE$:

$$\hbar \frac{dk}{dt} = -eE \quad \rightsquigarrow \quad k(t) = -\frac{eE}{\hbar}t \quad (4.82)$$

which will cause the states to move slightly to the side – but not indefinitely, presumably because of the impurities? There do exist cases where you can drive the band on and on towards the right. Once it disappears on the right side of the Brillouin zone boundary, it will reappear on the left side – these are known as Bloch Oscillations.

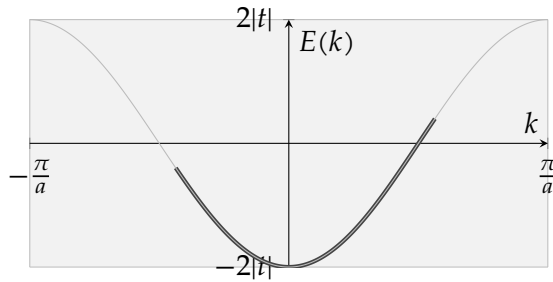


Figure 4.6: The first Brillouin zone for a monatomic, monovalent tight-binding chain. The electric field causes a displacement in k .

The group velocity $\frac{d\omega}{dk} = v_g$ is asymmetric now – there are more (quasi-) electrons with a positive group velocity than a negative group velocity – we have induced a current.

This leads to the fact that filled bands cannot carry currents – the band is periodic and filled, so an displacement will leave the band unchanged.

4.7 CRYSTAL STRUCTURE

A **lattice** is defined as an infinite set of points (lattice points) defined by integer linear combinations of a set of linearly independent primitive lattice vectors. In three dimensions we can write a lattice vector as

$$\mathbf{R}_{[n_1, n_2, n_3]} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 = [n_1, n_2, n_3], \quad (4.83)$$

$$n_1, n_2, n_3 \in \mathbb{Z} \quad (4.84)$$

Within the lattice is the **unit cell**: a region of space such that when many identical units are stacked together it tiles (completely fills) all of space and reconstructs the full structure. If a unit cell contains only one lattice point it is a **primitive unit cell**.

There is an algorithm that can be used to always create a primitive unit cell from lattice point. Given a lattice point, all points in space that are closer to the given lattice points than to any other lattice point constitute the **Wigner-Seitz cell**. Hence by finding perpendicular bisectors between the given and all of the closest lattice points you will find the Wigner-Seitz cell.

The unit cell describes the *lattice points*, which do not necessarily correspond to atoms. Hence given a unit cell, we need to define where the atoms lie within the cell – this is done using the **basis**. For instance graphene's basis is

$$\mathbf{R}_{[n_1 \ n_2]}^{\text{Carbon}_1} = [n_1, n_2] + \begin{bmatrix} 2 & 1 \\ 3 & 3 \end{bmatrix} \quad (4.85)$$

$$\mathbf{R}_{[n_1 \ n_2]}^{\text{Carbon}_2} = [n_1, n_2] + \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix} \quad (4.86)$$

Note that the basis is dependent on which unit cell is chosen – hence not unique.

This basis is for the following unit cell:

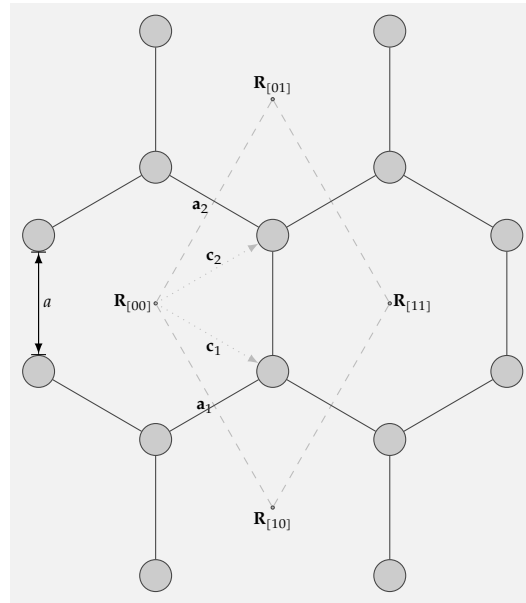


Figure 4.7: Graphene Lattice. The dashed lines denote the unit cell.

In three dimensions there are only a finite number of lattices, I will name two of these presently, as well as their defining characteristics:

1. **Body centred cubic (bcc) lattice:** A simple cubic lattice where we have added an additional lattice point in the middle of the cube:

$$\mathbf{R}_{\text{corner}} = [n_1, n_2, n_3] \quad (4.87)$$

$$\mathbf{R}_{\text{centre}} = [n_1, n_2, n_3] + \left[\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right] \quad (4.88)$$

with the following primitive unit vectors

$$\mathbf{a}_1 = [1, 0, 0] \quad (4.89)$$

$$\mathbf{a}_2 = [0, 1, 0] \quad (4.90)$$

$$\mathbf{a}_3 = \left[\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right] \quad (4.91)$$

2. **Face centred cubic (fcc) lattice:** A simple cubic lattice where we have added additional lattice points at the centre of each face:

$$\mathbf{R}_{\text{corner}} = [n_1, n_2, n_3] \quad (4.92)$$

$$\mathbf{R}_{\text{face}, xy} = [n_1, n_2, n_3] + \left[\frac{1}{2}, \frac{1}{2}, 0 \right] \quad (4.93)$$

$$\mathbf{R}_{\text{face}, xz} = [n_1, n_2, n_3] + \left[\frac{1}{2}, 0, \frac{1}{2} \right] \quad (4.94)$$

$$\mathbf{R}_{\text{face}, yz} = [n_1, n_2, n_3] + \left[0, \frac{1}{2}, \frac{1}{2} \right] \quad (4.95)$$

with primitive lattice vectors

$$\mathbf{a}_1 = \left[\frac{1}{2}, \frac{1}{2}, 0 \right] \quad (4.96)$$

$$\mathbf{a}_2 = \left[\frac{1}{2}, 0, \frac{1}{2} \right] \quad (4.97)$$

$$\mathbf{a}_3 = \left[0, \frac{1}{2}, \frac{1}{2} \right] \quad (4.98)$$

It can be quite difficult to show 3D crystal structures, therefore one often uses the **plan view**, where you project the 3D image onto the plane and explicitly denote the lost-component for each atom:

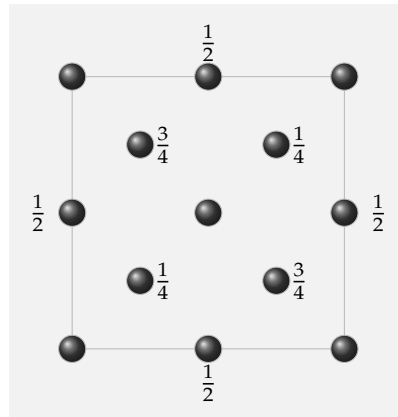


Figure 4.8: Plan view of the diamond crystal. Unmarked points are at $z = 0$ and $z = 1$.

4.8 THE RECIPROCAL LATTICE

Given a lattice vector, \mathbf{R} , we can define the reciprocal lattice as all vectors, \mathbf{G} , such that

$$e^{i\mathbf{G}\cdot\mathbf{R}} = 1 \quad (4.99)$$

We now make the following claims:

1. The reciprocal lattice is a lattice in reciprocal space.
2. The primitive reciprocal lattice vectors (\mathbf{b}_i) have the following property:

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij} \quad (4.100)$$

where \mathbf{a}_i are the primitive direct lattice vectors.

Using these properties, we can *construct* the vectors \mathbf{a}_i :

$$\mathbf{b}_i = \frac{\pi\epsilon_{ijk}\mathbf{a}_j \times \mathbf{a}_k}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad (4.101)$$

for instance, this gives us

$$\mathbf{b}_1 = \frac{2\pi\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad (4.102)$$

These vectors are linearly independent, and therefore they span \mathbb{R}^3 . Thus we can write all \mathbf{G} as *because \mathbf{a}_i are linearly independent*

$$\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3 \quad (4.103)$$

Additionally we have that

$$1 = e^{i\mathbf{G}\cdot\mathbf{R}} = e^{2\pi(n_1m_1+n_2m_2+n_3m_3)} \quad (4.104)$$

this is true if and only if $m_1, m_2, m_3 \in \mathbb{Z}$.

4.8.1 Fourier transform of the direct lattice

Let us define the density of lattice points as

$$\rho(\mathbf{r}) = \sum_{n_1, n_2, n_3} \delta^3(\mathbf{r} - (n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3)) \quad (4.105)$$

The Fourier transform hereof would be

$$\mathcal{F}\{\rho(\mathbf{r})\} = \sum_{n_1, n_2, n_3} \int_{\mathbb{R}^3} d\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} \delta^3(\mathbf{r} - (n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3)) \quad (4.106)$$

$$= \sum_{n_1, n_2, n_3} e^{i\mathbf{k}\cdot(n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3)} = \sum_{\mathbf{R} \in \text{DL}} e^{i\mathbf{k}\cdot\mathbf{R}} \quad (4.107)$$

$$= \frac{(2\pi)^3}{|\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|} \sum_{\mathbf{G}} \delta^3(\mathbf{k} - \mathbf{G}) \quad (4.108)$$

Where $|\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)| = V$ is the volume of the unit cell. However generally it is useful not to perform the integral over \mathbf{r} , and instead define the *structure factor*:

$$\mathcal{F}\{\rho(\mathbf{r})\} = \frac{(2\pi)^3}{V} \sum_{\mathbf{G}} \delta^3(\mathbf{k} - \mathbf{G}) S(\mathbf{k}) \quad (4.109)$$

where

$$S(\mathbf{k}) = \int_{\text{unit cell}} d\mathbf{x} e^{i\mathbf{k}\cdot\mathbf{x}} \rho(\mathbf{x}) \quad (4.110)$$

4.8.2 Reciprocal Lattice Points as Families of Lattice Planes

A *family of lattice planes* is an infinite set of equally separated parallel lattice planes, which taken together contain all points of the lattice.

Furthermore we claim that the family of planes are in one-to-one correspondence with the possible directions of reciprocal lattice vectors. The planes are normal to these reciprocal normal vectors. The spacing between the lattice planes is $d = 2\pi/\|\mathbf{G}_{\min}\|$, where \mathbf{G}_{\min} is the reciprocal lattice vector that is normal to the set of planes at has the shortest length.

Lattice Planes and Miller Indices

Let us begin by defining the Miller indices:

$$(h, k, \ell) \equiv h\mathbf{b}_1 + k\mathbf{b}_2 + \ell\mathbf{b}_3 = \mathbf{G}_{(h,k,\ell)} \quad (4.111)$$

additionally it is customary to write, for example $(1, \bar{1}, 1)$ instead of $(1, -1, 1)$. The shortest distance between planes:

$$d_{(h,k,\ell)} = \frac{2}{\|\mathbf{G}\|} = \frac{2\pi}{\sqrt{h^2\|\mathbf{b}_1\|^2 + k^2\|\mathbf{b}_2\|^2 + \ell^2\|\mathbf{b}_3\|^2}} \quad (4.112)$$

$$\rightsquigarrow \frac{1}{|d_{(h,k,\ell)}|^2} = \left(\frac{h}{a_1}\right)^2 + \left(\frac{k}{a_2}\right)^2 + \left(\frac{\ell}{a_3}\right)^2 \quad (4.113)$$

which is an ellipsoid in h, k and ℓ . Note that for cubic lattices ($a_1 = a_2 = a_3$) this simplifies to

$$d_{(h,k,\ell)}^{\text{cubic}} = \frac{a}{\sqrt{h^2 + k^2 + \ell^2}} \quad (4.114)$$

There is a point to make here:

1. For *unit cells* (h, k, ℓ) will describe a family of lattice planes, if and only if h, k and ℓ are (pairwise) coprime.
2. For non-primitive unit cells will generally not have h, k and ℓ to be pairwise coprime. For instance for a bcc crystal with the conventional cell $(0, 1, 0)$ does not represent a family of lattice planes, whereas $(0, 2, 0)$ does.

4.8.3 Brillouin Zones

The **Brillouin zone** is any primitive unit cell of the reciprocal lattice.

The first Brillouin zone is the set of \mathbf{k} points that are closer to $\mathbf{0}$ than any other reciprocal lattice point. Similarly the second Brillouin zone is the set of \mathbf{k} points where $\mathbf{0}$ is the next closest lattice point.

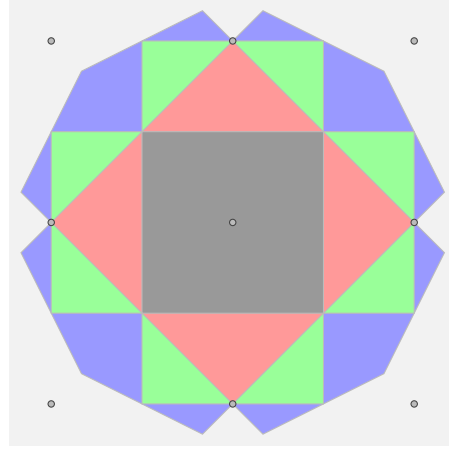


Figure 4.9: First four Brillouin zones for a square lattice. Gray: first Brillouin zone, red: second Brillouin zone, green: third Brillouin zone, lavender: fourth Brillouin zone.

4.9 WAVE SCATTERING BY CRYSTALS

According to Fermi's Golden Rule the scattering rate, $\Gamma(\mathbf{k}, \mathbf{k}')$, from $\mathbf{k} \rightsquigarrow \mathbf{k}'$ is given by:

$$\Gamma(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} |\langle \mathbf{k}' | V | \mathbf{k} \rangle|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) \quad (4.115)$$

Where

$$\langle \mathbf{k}' | V | \mathbf{k} \rangle = \frac{1}{L^3} \int d\mathbf{r} e^{-i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}} V(\mathbf{r}) \quad (4.116)$$

Now note that $V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r})$, for any lattice vector \mathbf{R} . Hence we can rewrite this as

$$\langle \mathbf{k}' | V | \mathbf{k} \rangle = \frac{1}{L^3} \left(\sum_{\mathbf{R}} e^{-i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{R}} \right) \left(\int_{\text{unit cell}} d\mathbf{r} e^{-i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}} V(\mathbf{r}) \right) \quad (4.117)$$

As before, the first term in parentheses must be zero unless

$$\mathbf{k}' - \mathbf{k} = \mathbf{G} \quad (4.118)$$

which is known as the **Laue condition**, which is equivalent to the conservation of crystal momentum, and hence $|\mathbf{k}'| = |\mathbf{k}|$. We use from before that the structure factor is given by

$$S(\mathbf{G}) = \int_{\text{unit cell}} d\mathbf{r} e^{i\mathbf{G} \cdot \mathbf{r}} V(\mathbf{r}) \quad (4.119)$$

The total potential is usually a linear combination of individual potentials:

$$V(\mathbf{r}) = \sum_{\text{atoms } j} V_j(\mathbf{r} - \mathbf{r}_j) \quad (4.120)$$

from where we get that

$$S(\mathbf{G}) = \sum_{\text{atoms } j \text{ in unit cell}} f_j(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}_j} \quad (4.121)$$

where $f_j(\mathbf{G})$ is the Fourier transform of the scattering potential for atom j :

$$f_j(\mathbf{G}) = \int d\mathbf{r} e^{i\mathbf{G} \cdot \mathbf{r}} V_j(\mathbf{r}) \quad (4.122)$$

4.10 ELECTRONS IN A PERIODIC POTENTIAL

4.10.1 Nearly Free Electron Model

A free electron has the following Hamilton and energy levels:

$$H_0 = \frac{\|\mathbf{p}\|^2}{2m}, \quad \varepsilon_0(\mathbf{k}) = \frac{\hbar^2 \|\mathbf{k}\|^2}{2m} \quad (4.123)$$

For periodic boundary conditions replace $\pi \rightsquigarrow 2\pi$

given the three quantisation lengths L_x, L_y and L_z we can write \mathbf{k} as

$$\mathbf{k} = \pi \left(\frac{n_x \mathbf{x}}{L_x} + \frac{n_y \mathbf{y}}{L_y} + \frac{n_z \mathbf{z}}{L_z} \right), \quad n_x, n_y, n_z \in \mathbb{Z} \quad (4.124)$$

Suppose now we perturb the free electron with a periodic potential:

$$H = H_0 + V(\mathbf{r}), \quad V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}) \quad (4.125)$$

First note the matrix element

$$\langle \mathbf{k}' | V | \mathbf{k} \rangle = \frac{1}{L^3} \int d\mathbf{r} e^{i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}} V(\mathbf{r}) \equiv V_{\mathbf{k}' - \mathbf{k}} \quad (4.126)$$

is zero unless $\mathbf{k}' - \mathbf{k} = \mathbf{G}$, a reciprocal lattice vector. This implies that the perturbation will only mix the state $|\mathbf{k}\rangle$ with $|\mathbf{k} + \mathbf{G}\rangle$. Thus, if we are far away from the Brillouin zone boundary we can use non-degenerate perturbation theory:

$$\varepsilon(\mathbf{k}) = \varepsilon_0(\mathbf{k}) + \langle \mathbf{k} | V | \mathbf{k} \rangle = \varepsilon_0(\mathbf{k}) + V_0 \quad (4.127)$$

henceforth we will assume that $V_0 = 0$. Second-order perturbation theory will thus give us

$$\varepsilon(\mathbf{k}) = \varepsilon_0(\mathbf{k}) + \sum_{\mathbf{k}' \neq \mathbf{k}} \frac{|\langle \mathbf{k}' | V | \mathbf{k} \rangle|^2}{\varepsilon_0(\mathbf{k}) - \varepsilon_0(\mathbf{k}')} \quad (4.128)$$

where \mathbf{k}' naturally is still restricted to be on the reciprocal lattice.

However, close to the Brillouin zone boundary we have that

$$\varepsilon_0(\mathbf{k}) \approx \varepsilon_0(\mathbf{k} + \mathbf{G}) \quad (4.129)$$

therefore we need to perform degenerate perturbation theory. Each state will only be degenerate with one other state – however, a state can be approximately degenerate with numerous other states, which would require us to increase the size of our degenerate space. However we will stick to two degenerate state, $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$. The perturbation matrix is given by

$$W = \begin{pmatrix} \langle \mathbf{k} | V | \mathbf{k} \rangle & \langle \mathbf{k}' | V | \mathbf{k} \rangle \\ \langle \mathbf{k} | V | \mathbf{k}' \rangle & \langle \mathbf{k}' | V | \mathbf{k}' \rangle \end{pmatrix} = \begin{pmatrix} \varepsilon_0(\mathbf{k}) & V_G \\ V_G^* & \varepsilon_0(\mathbf{k} + \mathbf{G}) \end{pmatrix} \quad (4.130)$$

The eigenvalues of the perturbation matrix are

$$E_{\pm} = \varepsilon_0(\mathbf{k}) \pm |V_G| \quad (4.131)$$

where we have assumed that $\varepsilon_0(\mathbf{k}) = \varepsilon_0(\mathbf{k} + \mathbf{G})$.

*that the energies are equal,
not approximately equal.*

Not quite at the boundary

This part will be done in one-dimension. If we are δ away from the Brillouin zone boundary, then it is the case that

$$\varepsilon_0(n\pi/a + \delta) \approx \varepsilon_0(-n\pi/a + \delta) \quad (4.132)$$

The unperturbed energies are

$$\varepsilon_0(\pm n\pi/a + \delta) = \frac{\hbar^2}{2m} \left(\left(\frac{n\pi}{a} \right)^2 \pm \frac{2n\pi\delta}{a} + \delta^2 \right) \quad (4.133)$$

and plugging these into the characteristic polynomial:

$$E_{\pm} = \frac{\hbar^2 n^2 \pi^2}{2ma^2} \pm |V_G| + \frac{\hbar^2 \delta^2}{2m} \left(1 \pm \frac{\hbar^2 n^2 \pi^2}{ma^2} \frac{1}{|V_G|} \right) \quad (4.134)$$

for small perturbations, compared to the eigenenergies of the free electron, which is exactly what we are looking at, the second term in the parentheses is greater than unity.

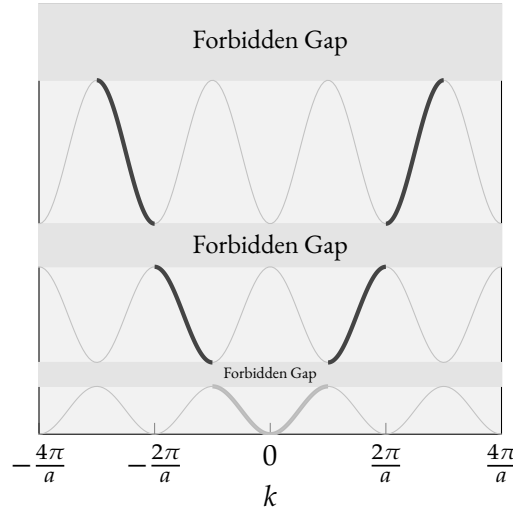


Figure 4.10: Dispersion of a nearly free electron model

Note the fact that the dispersion is parabolic close to the Brillouin zone boundary:

$$E_{\pm}(G + \delta) = C_{\pm} \pm \frac{\hbar^2 \delta^2}{2m_{\pm}^*}, \quad m_{\pm}^* = \frac{1}{|1 \pm \frac{\hbar^2 n^2 \pi^2}{2ma^2} \frac{1}{|V_G|}|} \quad (4.135)$$

4.10.2 Bloch's Theorem

The assumption that we can treat electrons as free particles, and hence describe their wave functions as planar waves seems rather bold. Especially because we previously spoke about scattering of electrons due to impurities. However **Bloch's theorem** states that the eigenfunctions of electrons in a periodic potential is of the form:

$$\Psi_{\mathbf{k}}^{\alpha}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u_{\mathbf{k}}^{\alpha}(\mathbf{r}) \quad (4.136)$$

where $u_{\mathbf{k}}^{\alpha}(\mathbf{r})$ is a periodic in the unit cell and \mathbf{k} can be chosen within the first Brillouin zone. Thus electrons behave as planar waves, with some periodic modulation of amplitude (and possibly phase, for $u_{\mathbf{k}}^{\alpha} \in \mathbb{C}$). In order for $u_{\mathbf{k}}^{\alpha}(\mathbf{r})$ it must be possible to write it as a sum over \mathbf{G} – every periodic function has a Fourier transform:

$$u_{\mathbf{k}}^{\alpha} = \sum_{\mathbf{G}} \tilde{u}_{\mathbf{G},\mathbf{k}}^{\alpha} e^{i\mathbf{G} \cdot \mathbf{r}} \quad (4.137)$$

hence an alternative expression of Bloch's theorem would be

$$\Psi_{\mathbf{k}}^{\alpha} = \sum_{\mathbf{G}} \tilde{u}_{\mathbf{G},\mathbf{k}}^{\alpha} e^{i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}} \quad (4.138)$$

4.II INSULATOR, SEMICONDUCTOR OR METAL

In short: insulators have filled bands, metals have partially filled bands. Semiconductors are a specific type of insulator, where the band gap is accessible through thermal fluctuations at room temperature. If the band gap is below 4eV the classify the insulator as a semiconductor.

4.II.1 Energy bands in two and three dimensions

If the unit cell is square, then the reciprocal unit cell is also square – hence the Brillouin zone is square. If we now fill up "empty space" with electrons from $E = 0$ to $E = E_F$ we will first fill low energy states, and then slowly build upwards once all the states filled, resulting in a circular Fermi-"surface". If we now add a strong periodic potential, we will reduce the band-energies close to the inside of the first Brillouin zone boundary, as before, and increase the energies close to the outside of the first Brillouin zone.

Monovalent atoms

For vary potential strength in a monovalent lattice we get:



Figure 4.11: Left: free electrons. Middle: intermediate potential. Right: strong potential

Divalent atoms

For varying potential strengths in a divalent lattice we get:



Figure 4.12: Left: free electrons. Right: electrons in a very strong periodic potential

Every periodic function with only a finite number of discontinuities per period, which has a bounded variation and must be absolutely integrable over a period.

The image that only states below the fermi energy are occupied is only true at $T = 0$

in 2D the Fermi-surface is a line.

For any potential in the spectrum between no potential and strong potential, there will be partial filling of the second Brillouin zone:

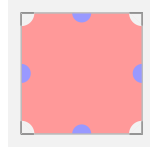


Figure 4.13: *Electrons in a square lattice with intermediate potential strength. The red shaded area is the occupied part of the first Brillouin zone, and the blue shaded area is the occupied part of the second Brillouin zone*

Here the same principle holds as in 1D: filled bands do not conduct electricity – for example a divalent square lattice will not conduct electricity. However, now it is possible for divalent square lattices to conduct electricity, providing the potential is weak enough.

In three dimensions the Fermi surface is a surface, which behaves similarly to what we discussed now. Weak potential have spherical Fermi surfaces. Intermediate potentials will have 6 bulges, which touch the Brillouin zone boundary for strong potentials.

4.II.2 *Failures of the Band-structure picture of metals*

We have discussed how to determine the electric properties of solids, by first calculating the band structure, which depends on the geometry of the solid, and then filling the band up with however many electrons we have, following Pauli's exclusion principle. From there we had the simple principle: filled bands are inert.

However, this would cause us to assume that all monovalent lattices are metals, however, we have completely ignored the interaction between electrons, which does play a significant role in some systems. An example of this picture failing is in magnetism: the magnetic moments of electrons can interact – if they all align you have a ferromagnet, which shows complexity that simply cannot be described by our model.

Another example is the Mott insulator: if the electronic repulsion is so strong in a monovalent lattice that it simply costs too much to have two electrons at the same atom (irrespective of spin), then the material will be inert, even though it is monovalent.

4.II.3 *Optical properties of metals and insulators*

If a semiconductor has a *direct* band gap ΔE , then a photon with this energy can excite an electron from the valence to the conduction band. Conversely an electron in the conduction band can de-excite into the valence band, giving off a photon with the corresponding energy. Thus if a material has a band gap in the visible region of the electromagnetic spectrum, this material will absorb light.

If a material has a band gap greater than $\sim 3.2\text{eV}$ then light will not interact: the material is transparent. For example this is the case with diamond and quartz.

Cadmium sulfide has a band gap of 2.6eV which corresponds to green-blue light. Thus *CdS* appears reddish (removing green-blue and upwards from white gives you red).

Metals conduct electricity, and therefore their interaction with light is slightly different. The electronic excitation is quickly undone, such that most absorbed photons are re-emitted quickly (hence reflected). If all photons are re-emitted the metal will look silver, like a mirror. However, some metals do absorb some frequencies, depending on the band structure and potential strength, giving them colours – normally yellow-ish or red-ish.

This is an oversimplified model of metals.

4.12 SEMICONDUCTOR PHYSICS

4.12.1 Holes

Relativistic quantum mechanics teaches us to think about negative energy states. And to all of our surprises one often comes across relativistic quantum mechanical equations in condensed matter physics. Hence it is natural to think about these negative energy states, in the context of condensed matter physics.

Imagine we have a filled valence band, and an empty conduction band. Now we excite a single electron, leaving a *hole* in its place. Now we can either do our calculations on the electron in the conduction band, or on the hole in the valence band. We will get the same result.

The energy is not exactly negative in this context though – if you set the zero value of energy to be exactly the middle between the gaps then it is clear that the hole will have the negative of the electrons energy.

In fact, the negative energy states I mentioned previously are the holes. It can be very convenient to perform calculations using holes instead of electrons, and in fact in some materials it is the holes and not the electrons that carry current, which is why we had Hall-coefficients with the wrong sign previously.

4.12.2 Effective Mass of Electrons and Holes

The dispersion relation for an electron near the bottom of the conduction band can be Taylor expanded:

$$E = E_{\min}^C + \alpha \|\mathbf{k} - \mathbf{k}_{\min}\|^2 + \dots \quad (4.139)$$

which, as we discussed previously leads us to define the effective mass:

$$m_e^* = \frac{\hbar^2}{2\alpha} \quad (4.140)$$

and equivalently we can define the group velocity:

$$\mathbf{v} = \frac{\nabla_{\mathbf{k}} E}{\hbar} = \frac{\hbar(\mathbf{k} - \mathbf{k}_{\min})}{\hbar} \quad (4.141)$$

Similarly the energy of the hole can be Taylor expanded about the maximum of the valence band:

$$E = E_{\max}^V - \alpha \|\mathbf{k} - \mathbf{k}_{\min}\|^2 + \dots \quad (4.142)$$

From here it is conventional to define the hole mass to be positive, which has some implications later on, when we discuss hole momentum and velocity. Thus

$$\frac{\hbar}{m_h^*} = -\frac{\partial^2 E}{\partial k^2} \rightsquigarrow m_h^* = \frac{\hbar^2}{2\alpha} \quad (4.143)$$

It requires positive energy to push an electron *upwards* in the band. Viewing a hole as the lack of an electron will thus lead us to the conclusion that it requires positive energy to push a hole *downwards* in the valence band. This is rather counterintuitive, but is in fact the case. Hence holes prefer to maximise their energy, whereas electrons minimise theirs: The minimal energy state is an electron at the bottom of the conduction band and a hole at the top of the valence band.

4.12.3 Momentum of holes

If we remove an electron with momentum $\hbar\mathbf{k}$ from a filled valence band the valence band will have momentum $-\hbar\mathbf{k}$ —remember filled bands carry no momentum. Thus it is convenient to define

$$\mathbf{k}_{\text{hole}} = -\mathbf{k}_{\text{electron}} \quad (4.144)$$

Thus

$$\mathbf{v}_{\text{hole}} = \mathbf{v}_{\text{missing electron}} \quad (4.145)$$

Where we have used that

$$E_{\text{absence of electron}}(\mathbf{k}) = -E_{\text{hole}}(\mathbf{k}) \quad (4.146)$$

This is natural, because if we remove an electron then we decrease the energy by the energy that the electron had.

4.12.4 Adding electrons or holes: Doping

Imagine a lattice of Silicon atoms, with an occasional Phosphorous atom. If the concentration of Phosphorous is negligible compared to that of Silicon, then the impurities will not have a measurable effect on the dispersion relation. However, Phosphorous has an additional electron (and proton), which means that there is a free electron floating around – this is called *n*-doping, because we dope with negative charges.

Similarly, if we dope the Silicon with Aluminium we will have one fewer electron per Aluminium atom: we have ”added” holes. This is *p*-doping, because we dope with positive charges.

Impurity states

The additional charges, both electrons and holes, only see the impurities – the lattice due to its periodicity is effectively invisible. However, the lattice does affect the dielectric constant ($\epsilon_0 \rightsquigarrow \epsilon$ where $\epsilon = \epsilon_0 \epsilon_r$), which in turn affects the energy levels and characteristic lengths. Additionally it is not electrons that are bound by the impurities, but the quasi-particles made by the hybridised wave-functions: we need to take the effective mass into account.

For low concentrations of dopants (which is the limit we require for our model to work anyway) we can treat the impurities as hydrogenic atoms, with a relative permittivity. We get

$$Ry^{\text{eff}} = Ry \left(\frac{m_e^*}{m} \frac{1}{\epsilon_r^2} \right), \quad Ry \approx 13.6 \text{ eV} \quad (4.147)$$

$$a_0^{\text{eff}} = a_0 \left(\epsilon_r \frac{m}{m_e^*} \right), \quad a_0 \approx 0.5 \text{ \AA} \quad (4.148)$$

The effective Rydberg constant is usually much smaller than the band gap.

For typical material Ry^{eff} is below 0.1 eV and a_0 is above 30 \AA.

These hydrogenic states are added to the band structure. If the material is n -doped the impurity band is just below the bottom of the conduction band (donor impurity states), for p -doped materials the impurity band is just above the top of the valence band (acceptor impurity state).

4.12.5 Statistical Physics of Electrons and Holes

If the chemical potential is "well" below the conduction band ($\beta(\epsilon_c - \mu) \gg 1$) we can approximate Fermi-Dirac statistics by Boltzmann statistics, which gives us the standard expression for the total density of electrons in the conduction band:

$$n(T) = \frac{1}{4} \left(\frac{2m_e^* k_B T}{\pi \hbar^2} \right)^{\frac{3}{2}} e^{-\beta(\epsilon_c - \mu)} \quad (4.149)$$

We used the density of states of the conduction band in the calculation ($n(T) = \int_{\epsilon_c}^{\infty} d\epsilon g_c(\epsilon) n_F(\beta(\epsilon - \mu))$). When we calculate the total density of holes in the valence band we need to use the density of states of the valence band and replace the Fermi-Dirac distribution by $1 - n_F$ – the states not occupied by electrons are occupied by holes. From here we get

$$p(T) = \frac{1}{4} \left(\frac{2m_h^* k_B T}{\pi \hbar^2} \right)^{\frac{3}{2}} e^{-\beta(\mu - \epsilon_v)} \quad (4.150)$$

the two different exponents are due to the limits of integration (the conduction band is integrated from ϵ_c to infinity, whereas the valence band is integrated from $-\infty$ to ϵ_v). An important quantity is

$$n(T)p(T) = \frac{(m_h^* m_e^*)^{\frac{3}{2}}}{16} \left(\frac{2k_B T}{\pi \hbar^2} \right)^3 e^{-\beta E_{\text{gap}}}, \quad E_{\text{gap}} = \epsilon_c - \epsilon_v \quad (4.151)$$

This is the *law of mass action*. In order to see that this is a law of mass action, consider an intrinsic semiconductor (the number of holes equals the number of electrons): Now think of the following reaction:



where p^+ is the hole and $[e^-] = n(T)$ and $[p^+] = p(T)$. The concentration $[e^- p^+]^0 = 1$, so

$$\frac{[e^-][p^+]}{[e^- p^+]^0} = n(T)p(T) = K(T) \quad (4.153)$$

Intrinsic semiconductors

As mentioned before $n = p$ for intrinsic semiconductors, hence by diving our expressions for $n(T)$ and $p(T)$ and taking the logarithm we can solve for $\mu(T)$:

$$\mu(T) = \frac{1}{2} (\epsilon_c + \epsilon_v) + \frac{3}{4} k_B T \ln \left(\frac{m_h^*}{m_e^*} \right) \quad (4.154)$$

At $T = 0$ the chemical potential lies exactly between the extrema of the conduction and valence bands.

Interestingly enough for $m_h^* = m_e^*$, which means the shape of the bottom of the conduction is the same as the top of the valence band, there is no temperature dependence of the chemical potential.

Doped semiconductors

We can define the doping, D :

$$D = n - p \quad (4.155)$$

Negatively doping a material will begin to fill the conduction band from the bottom, which results in the chemical potential rising up towards the conduction band. Similarly positively doping a material will increase the number of holes in the valence band, which in turn causes the chemical potential to fall towards the top of the valence band.

4.13 SEMICONDUCTOR DEVICES

4.13.1 *Designing Band Gaps*

The band gaps of $GaAs$ and $AlAs$ are 1.4eV and 2.7eV respectively. Both have direct band gaps at $\mathbf{k} = \mathbf{0}$. If we now were to mix these two materials to $Al_xGa_{1-x}As$ one could hope that the band gap is given by the convex combination:

$$E_{\text{gap}}(x) = (1 - x)1.4\text{eV} + x \cdot 2.7\text{eV}, \quad x \in (0, 0.4) \quad (4.156)$$

This turns out to be approximately true, for x between 0 and 0.4. Therefore by choosing the relative concentrations we can engineer a semiconductor with a band gap anywhere in the interval (1.4eV, 1.92eV).

Using this we can create a sandwich with $Al_xGa_{1-x}As$ on each side of $GaAs$. The energy gap will have an abrupt dip, in which one can fill with bound electron states. Effectively you can use this method to make a finite well potential. The importance is that electrons in the conduction band see the bottom of the band as a potential as a function of position, which they can be trapped in, or scattered by.

4.13.2 *pnunction*

Putting an n -doped material into contact with a p -doped material has some very interesting effects. The n -doped material has:

1. A filled valence band

2. A mostly empty conduction band

whereas a p -doped material has:

1. A mostly filled valence band (mostly empty if you're thinking about holes)
2. An empty conduction band (full if you're thinking about holes)

When these are brought into contact the electrons in the n -doped material's conduction band will begin to annihilate the holes in the p -doped material's valence band. This results in a depletion zone – a zone where all the holes and electrons have been annihilated. In this depletion zone there will be an electric field, pointing from n -doped towards p -doped. This electric field obstructs further electrons to travel through the depletion zone (and similarly prevents holes to do so). The electric field is such that the electrochemical potential ($\mu - e\phi$) is equal on each side of the depletion zone, where ϕ is the electric potential.

Solar Cell

If one applies light to a p - n junction electron-hole pairs may appear (if the photon energy is greater than the band-gap). Once this pair is created, and if they have enough kinetic energy to avoid each other initially, they will be accelerated in opposite directions by the electric field in the depletion zone: electrons are sent towards the n -doped region whereas holes are sent towards the p -doped region. This results in a current going from n to p . This is, in principle, how photovoltaics work.

The Diode

By applying an additional, external, electric potential (voltage) to a p - n junction, and by applying it in the opposite direction of the electric field in the depletion zone one can get electrons (holes) to pass through the depletion zone and enter the p - (n -) doped material. Once they are there they will annihilate with holes (electrons). This results in a net current from p to n . The probability of this event occurring is proportional to $e^{-(E_{\text{gap}} + eV)/k_B T}$, where V is the applied voltage.

It is also possible for another effect to occur: an electron in the p -doped side is thermally excited into the conduction band. Once it is here it is free to move and the electric field in the depletion zone will accelerate it to the n -doped side. This however requires electrons (and holes) to be thermally excited, which has a probability that is proportional to $e^{-E_{\text{gap}}/k_B T}$.

Combining the two effects we get

$$I_{\text{total}} = J_s(T) (e^{-eV/k_B T} - 1) \quad (4.157)$$

where $J_s \propto e^{-E_{\text{gap}}/k_B T}$. This is known as the "Diode equation" and says that current flows far more easily from p to n , than the other way resulting in a very important electric component.

4.13.3 *Transistor*

A transistor is a component with three terminals (source, drain and gate). You can build a transistor by putting an n -doped semiconductor between two p -doped semiconductors, or vice versa. The former is called a p -MOSFET, and the latter a n -MOSFET. We will focus on the n -MOSFET:

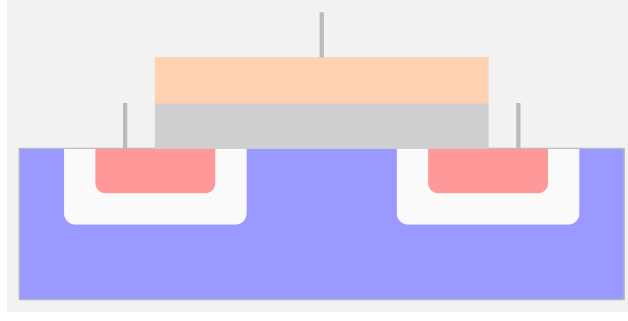


Figure 4.14: *n*-MOSFET. Red: *n*-doped semiconductor. Blue: *p*-doped semiconductor. White: depletion zone. Gray: Oxide insulator. Orange: Metal Gate. The two *n*-doped semiconductors are the source and drain terminals.

Applying a positive voltage to the metal gate will push away positive charges in the top of the *p*-doped semiconductor. Once $V > V_{\text{threshold}}$ this will breach a gap through which current can flow:

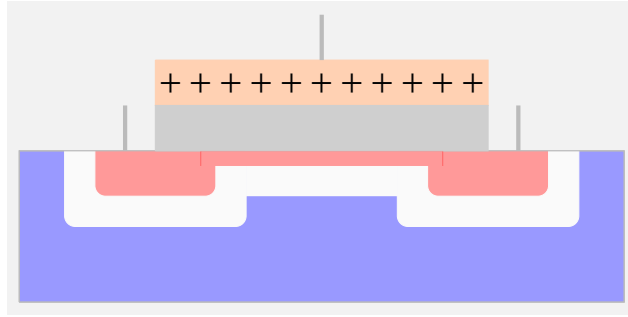


Figure 4.15: *n*-MOSFET with positive voltage at the metal gate. Red: *n*-doped semiconductor. Blue: *p*-doped semiconductor. White: depletion zone. Gray: Oxide insulator. Orange: Metal Gate. The two *n*-doped semiconductors are the source and drain terminals.

And that is how a *n*-MOSFET works. Oh, and MOSFET stands for **M**etal on top of **O**xide Insulator on top of a **S**emiconductor- **F**ield **E**ffect **T**ransistor.

Remember:

$$N = V \int_0^\infty d\varepsilon g(\varepsilon) n_F(\beta(\varepsilon - \mu)) \quad (4.158)$$

$$N = \frac{2L^d}{(2\pi)^d} \int d^d \mathbf{k} n_F(\beta(\varepsilon(\mathbf{k}) - \mu)) \quad (4.159)$$

$$\langle E \rangle = V \int_0^\infty d\varepsilon \varepsilon g(\varepsilon) n_F(\beta(\varepsilon - \mu)) \quad (4.160)$$

$$N = \frac{2L^d}{(2\pi)^d} \int d^d \mathbf{k} \varepsilon(\mathbf{k}) n_F(\beta(\varepsilon(\mathbf{k}) - \mu)) \quad (4.161)$$

5.1 ATOMIC PHYSICS

The magnetic moment of an infinitesimal current loop is given by

$$d\boldsymbol{\mu} = I d\mathbf{S} \quad \rightsquigarrow \quad \boldsymbol{\mu} = I \int_{\sigma} d\mathbf{S} \quad (5.1)$$

where I is the current and $d\mathbf{S}$ is the area. Integrating over a surface σ gives us the total magnetic moment.

Magnetic moments of atoms

Atoms have magnetic moments, which can be decomposed into three constituents: A nuclear moment, and two electronic moments, one due to its spin and one due to its orbital momentum:

$$\boldsymbol{\mu} = \boldsymbol{\mu}_N + \boldsymbol{\mu}_S + \boldsymbol{\mu}_L = \mu_N g_N \mathbf{S}_N + \mu_B g_s \mathbf{S} + \gamma \mathbf{L}, \quad \mu_B = \frac{e\hbar}{2m_e} \quad (5.2)$$

The orbital moment is demonstrated by the Einstein-de Haas effect, where the current through a coil with a ferromagnetic core is abruptly stopped causing slight a rotating of the core.

Classically applying a magnetic field to an orbital magnetic moment will cause a gyromagnetic precession, with frequency γB :

$$\text{torque} \equiv \mathbf{G} = \boldsymbol{\mu} \times \mathbf{B} = \frac{1}{\gamma} \dot{\boldsymbol{\mu}} \quad (5.3)$$

$$\rightsquigarrow \quad \boldsymbol{\mu}(t) = \boldsymbol{\mu} \cos(\Omega t + \phi_0) + \frac{\boldsymbol{\mu} \times \mathbf{B}}{B \cos(\theta)} \sin(\Omega t + \phi_0) \quad (5.4)$$

where $\Omega = \gamma B$ is the Larmor precession frequency and θ is the angle between $\boldsymbol{\mu}$ and \mathbf{B} .

Magnetisation and field

The flux density is simply a linear combination of the applied field and the magnetisation:

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad (5.5)$$

where

$$\mathbf{M} = \chi \mathbf{H}, \quad M_i = \chi_{ij} H_j \quad (5.6)$$

where the former is for linear materials where $\chi_{ij} = M \delta_{ij}$ and the latter is for non-linear materials: in non-linear materials the magnetisation can point in a different direction than the applied field. For linear materials we have

$$\mathbf{B} = \mu_0 (1 + \chi) \mathbf{H} = \mu_0 \mu_r \mathbf{H} \quad (5.7)$$

This course primarily uses the susceptibility χ rather than the relative permeability, μ_r .

The Bohr-van Leeuwen theorem

The total energy of a classical electron in a magnetic field is given by

$$E = \frac{\|\mathbf{p} + e\mathbf{A}\|^2}{2m} \quad (5.8)$$

This means that the partition function can be calculated

$$Z \sim \iint \cdots \int \exp(-\beta E(\mathbf{r}_i, \mathbf{p}_i)) d\mathbf{r}_1 \cdots d\mathbf{r}_N d\mathbf{p}_1 \cdots d\mathbf{p}_N \quad (5.9)$$

however as the momentum integral goes over entire \mathbb{R} we can shift the momentum by $e\mathbf{A}$ resulting in an integral that is independent of \mathbf{A} and hence all physical quantities are independent of \mathbf{A} – there can exist no global magnetism, classically: this is the statement of the Bohr-van Leeuwen theorem.

Quantum mechanics of spin

The z-axis projection of the spin of a particle is measured using

$$s_z |s m_s\rangle = \hbar m_s |s m_s\rangle \quad (5.10)$$

The factors of \hbar is often omitted. However, the x and y components are slightly more complicated, as $|s\rangle$ is not an eigenket to these operators. However, to make things easier we can define raising and lowering operators:

$$s_+ = \frac{1}{2} (s_x + i s_y) \quad (5.11)$$

$$s_- = \frac{1}{2} (s_x - i s_y) \quad (5.12)$$

Which means that

$$s_x = \frac{1}{2} (s_+ + s_-) \quad (5.13)$$

$$s_y = \frac{1}{2i} (s_+ - s_-) \quad (5.14)$$

which allows us to calculate what $\mathbf{s}|s m_s\rangle$ is:

$$\mathbf{s}|s m_s\rangle = \begin{pmatrix} \frac{\hbar}{2} (A_- |s m_s + 1\rangle + A_+ |s m_s - 1\rangle) \\ \frac{\hbar}{2i} (A_- |s m_s + 1\rangle - A_+ |s m_s - 1\rangle) \\ \hbar m_s |s m_s\rangle \end{pmatrix} \quad (5.15)$$

Where

$$A_{\pm} = \sqrt{(s \mp m_s)(s \pm m_s + 1)} \quad (5.16)$$

Coupling of two spins

Let us consider two spin- $\frac{1}{2}$ particles that are coupled through a Hamiltonian:

$$\hat{H} = A \mathbf{S}_1 \cdot \mathbf{S}_2 = \sum_i A S_1^i \otimes S_2^i \quad (5.17)$$

we use the direct product, because the two spins live in each their Hilbert spaces, and the coupling of the operators couples the Hilbert space. We know that

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1 \quad (5.18)$$

which implies that this system splits in to a singlet (of spin zero) and a triplet (with spin 1). This can also be seen by rewriting $\mathbf{S}_1 \cdot \mathbf{S}_2$ out with conserved quantities:

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2} (S^2 - S_1^2 - S_2^2), \quad \mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \quad (5.19)$$

where $S \in \{0, 1\}$. In conclusion

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \begin{cases} \frac{1}{4} & \text{triplets : } |\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ -\frac{3}{4} & \text{singlet : } \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \end{cases} \quad (5.20)$$

Generally

$$S_1 \otimes S_2 = |S_1 - S_2| \oplus (|S_1 - S_2| + 1) \oplus \dots \oplus (S_1 + S_2) \quad (5.21)$$

5.2 ISOLATED MAGNETIC MOMENTS

By beginning with the Hamiltonian for an electron:

$$\hat{H} = \sum_i \left(\frac{\|\boldsymbol{\sigma} \cdot (\mathbf{p}_i + e\mathbf{A}(\mathbf{r}_i))\|^2}{2m_e} + V_i \right) \quad (5.22)$$

you can show that

$$\hat{H} = \hat{H}_0 + \mu_B (\mathbf{L} + g_s \mathbf{S}) \cdot \mathbf{B} + \frac{e^2}{8m_e} \sum_i (\mathbf{B} \times \mathbf{r}_i)^2 \quad (5.23)$$

in the symmetric gauge, where

$$\hat{H}_0 = \sum_i \left(\frac{p_i^2}{2m_e} + V_i \right) \quad (5.24)$$

The first term that is added to \hat{H}_0 is the paramagnetic term, and the second is the diamagnetic term

Diamagnetism

The diamagnetic term has an expectation value of

$$E_{\text{dia}} = \frac{e^2 B^2}{8m} \sum_i \langle x_i^2 + y_i^2 \rangle = \frac{e^2 B^2}{12m} \sum_i \langle r_i^2 \rangle \quad (5.25)$$

where the final equality is only true if we assume spherical symmetry. The magnetisation is then given by

$$M = -\frac{\partial F}{\partial B} = -\frac{N}{V} \frac{\partial E_{\text{dia}}}{\partial B} = -\frac{Ne^2 B}{6m_e V} \sum_i \langle r_i^2 \rangle \quad (5.26)$$

For weak fields (or small susceptibilities) we can use $\chi = \frac{\partial M}{\partial H} \approx \mu_0 \frac{\partial M}{\partial B}$:

$$\chi = -\frac{N e^2 \mu_0}{V 6m_e} \sum_i \langle r_i \rangle \quad (5.27)$$

the sum $\sum_i \langle r_i \rangle$ goes over all electrons in the atomic shell.

Paramagnetism

As we have seen, the magnetic moment of an atom is, primarily, associated with the total orbital angular momentum of the electrons:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (5.28)$$

The partition function for the quantum mechanical description is

$$Z = \sum_{m_J=-J}^J \exp\left(\frac{m_J g_J \mu_B B}{k_B T}\right) \quad (5.29)$$

which gives us $\langle m_J \rangle$ under thermodynamic equilibrium, and hence

$$M = n g_J \mu_B \langle m_J \rangle = n k_B T \frac{\partial \ln Z}{\partial B} \quad (5.30)$$

Using the geometric series and definition the *Brillouin function*:

$$B_J(y) = \frac{2J+1}{2J} \coth\left(\frac{(2J+1)y}{2J}\right) - \frac{1}{2J} \coth\left(\frac{y}{2J}\right) \quad (5.31)$$

as well as the saturation magnetisation:

$$M_s = n g_J \mu_B J \quad (5.32)$$

it can be shown that

$$M = M_s B_J\left(\frac{g_J \mu_B J B}{k_B T}\right) \quad (5.33)$$

It is noteworthy that $B_{\frac{1}{2}}(y) = \tanh(y)$ and $B_{\infty}(y) = L(y)$, where

$$L(y) = \coth(y) - \frac{1}{y} \quad (5.34)$$

is the *Langevin function*, which is the semiclassical result.

For weak fields, where we can treat the Brillouin as linear:

$$B_J(y) = \frac{(J+1)y}{3J} + O(y^3) \quad (5.35)$$

we can calculate the susceptibility:

$$\chi = \frac{M}{H} \approx \frac{\mu_0 M}{B} = \frac{n \mu_0 \mu_{\text{eff}}^2}{3 k_B T}, \quad \mu_{\text{eff}} = g_J \mu_B \sqrt{J(J+1)} \quad (5.36)$$

and

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \quad (5.37)$$

The quantity $k_B T \mu_B = 460 \text{ T}$, we're essentially always in the weak field limit.

We see that this model is in agreement with the Curie law, which predicts that $\chi \sim T^{-1}$.

Hund's rules and the ionic ground state

The non-relativistic limit of the Dirac equation gives us, among other terms, a term that couples the angular momentum to the spin of an electron: $\mathbf{L} \cdot \mathbf{S}$. This interaction leads to the fine structure of atoms. We treat this exactly how we treated the coupling of two spins:

$$\lambda \langle \mathbf{L} \cdot \mathbf{S} \rangle = \frac{\lambda}{2} (J(J+1) - S(S+1) - L(L+1)) \quad (5.38)$$

where $J = |L - S|, |L - S| + 1, \dots, L + S \stackrel{S=\frac{1}{2}}{=} L \pm \frac{1}{2}$. This leads to the *Landé interval rule*:

$$E_J - E_{J-1} = \lambda J \quad (5.39)$$

We see that depending on the sign of λ the ground state will either have $J_{\text{gs}} = \max \{J\}$ or $J_{\text{gs}} = \min \{J\}$: this is exactly what Hund's third rule says.

Hund's rule

For *equivalent electrons* the ground state is found by:

1. Maximising the spin-multiplicity and hence S of the shell.
2. Maximising the angular momentum of the shell
3. Minimising J if the shell is under half-filled and otherwise maximising J .

These rules do not always work, but usually give a good indication of what the ground state is.

For instance Dy^{3+} has an outer shell of $4f^9$. $f = 3$ therefore the Pauli exclusion principle doesn't allow for more than 7 spin up electrons, therefore $S = \frac{5}{2}$. The seven spin up electrons also have $L = 0$ therefore we maximise L by letting the two spin down electrons have $m_L = 3$ and $m_L = 2$, so $L = 5 = H$. The shell is over half-full so we maximise $J = \frac{15}{2}$:

$$5f^9 \rightsquigarrow {}^6H_{\frac{15}{2}} \quad (5.40)$$

This only holds in the LS -coupling scheme, where S, L, J are good quantum numbers. For jj -coupling we would refer to levels using $(j_1, j_1)_J$ tuples. The LS -coupling scheme is valid when the spin-orbit coupling can be treated as a perturbation to the residual electrostatic interaction.

5.3 CRYSTAL FIELDS

In order to describe the magnetic properties of crystals it is necessary to investigate how magnetic moments interact with their surroundings. Particularly when the surrounding atoms are arranged in a crystal structure.

We will consider a d -orbital both an octahedral and tetrahedral crystal structure. The former has atoms placed on the axes with respect to the atom we are looking at, which means that the d_{z^2} and $d_{x^2-y^2}$ orbitals overlap (these point along the axes), whereas the remaining 3 orbitals do not overlap. We denote the orbitals that point along the axes as e_g and those that don't as t_{2g} . The energy

splitting between e_g and t_{2g} will be given by some energy Δ , and $E(e_g) - E_{\text{free}} = \frac{3\Delta}{5}$ whereas $E(t_{2g}) - E_{\text{free}} = -\frac{2\Delta}{5}$.

The tetrahedral splitting is similar, however now it is the t_{2g} that overlap, and hence $E(t_{2g}) - E_{\text{free}} = \frac{2\Delta}{5}$, whereas $E(e_g) - E_{\text{free}} = -\frac{3\Delta}{5}$.

Orbital Quenching

For strong crystal fields (stronger than the spin-orbit interaction) the half-filled and filled orbitals have $L = 0$, this effect is due to orbital quenching. Assuming the ground state is non-degenerate, which is easily realised by the crystal-field splitting it is also a real function. \hat{L} is an imaginary operator therefore $\langle \hat{L} \rangle = 0$.

The Jahn-Teller effect

The magnetic properties of a material can affect the structure of the material. A distortion of the crystal structure can decrease/increase the overlap of orbitals and therefore lower/increase the energy. Let us assume that this energy is linear in the distortion, and furthermore assume that the energy required to distort the crystal is quadratic in the distortion. This gives:

$$E(Q) = \pm AQ + \frac{1}{2}M\omega^2 Q^2 \quad \rightsquigarrow \quad Q_{\min} = \frac{\pm A}{M\omega^2} \quad (5.41)$$

This simple analysis implies that a distortion can cause the crystal structure to change, such that the energy goes from one local minimum to another: this is the Jahn-Teller effect.

5.4 INTERACTIONS

We would like to explain the long range order of ferro- and anti-ferromagnets, and will look at two interactions: the magnetic dipolar interaction, which is too small to be relevant at room temperature and the exchange interaction which is what causes magnetic ordering.

Magnetic dipolar interaction

The energy of the interaction between two magnetic dipoles is given by

$$E = \frac{\mu_0}{4\pi r^3} \left(\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2 - \frac{3}{r^2} (\boldsymbol{\mu}_1 \cdot \mathbf{r})(\boldsymbol{\mu}_2 \cdot \mathbf{r}) \right) \quad (5.42)$$

which for $r \approx 1\text{\AA}$ is $E \sim 60\text{eV}$ and hence far too small to result in macroscopic ordering at room temperature. Note that $k_B T \approx 26\text{meV}$ at room temperature which corresponds to the phononic energy.

Exchange interaction

Let us consider two identical atoms, that interact through the following Hamiltonian:

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{H}_{\text{int}} \quad (5.43)$$

where $\hat{H}_i|i\rangle = E_i|i\rangle$ and \hat{H}_{int} is the interaction between the two atoms. This interaction can be treated with first order perturbation theory, which first of all gives us new wavefunctions:

$$|\psi_S\rangle\chi_A = \frac{1}{\sqrt{2}} (|\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\rangle + |\psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1)\rangle)\chi_A \quad (5.44)$$

$$|\psi_A\rangle\chi_S = \frac{1}{\sqrt{2}} (|\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\rangle - |\psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1)\rangle)\chi_S \quad (5.45)$$

where I have included that spin-wave function that are required to make the total wave function anti-symmetric under particle exchange, because electrons are fermions. This gives us two energies: the two isolated atoms split into two states, one where the electrons are in the triplet ordering (χ_S) and one where the electrons are in the singlet ordering (χ_S), *just* like what we saw with $\mathbf{S}_1 \cdot \mathbf{S}_2$. Therefore, instead of writing a whole lot of integrals and complicated formulae for the exchange interaction, let us use $\mathbf{S}_1 \cdot \mathbf{S}_2$ and some multiplicative constant that is calculated with the integrals:

$$\hat{H}_{\text{int}} = -2JS_1 \cdot S_2 \quad (5.46)$$

where

$$J = \frac{V - SU}{1 - S^2} \quad (5.47)$$

$$U \equiv \int d\mathbf{r}_1 d\mathbf{r}_2 \hat{H}_{\text{int}} |\phi_1(\mathbf{r}_1)|^2 |\phi_2(\mathbf{r}_2)|^2 \quad (5.48)$$

$$V \equiv \int d\mathbf{r}_1 d\mathbf{r}_2 \hat{H}_{\text{int}} \psi_1^*(\mathbf{r}_2) \psi_2^*(\mathbf{r}_1) \psi_1(\mathbf{r}_1) \psi_2(\mathbf{r}_2) \quad (5.49)$$

$$S \equiv \int d\mathbf{r} \psi_1^*(\mathbf{r}) \psi_2(\mathbf{r}) \quad (5.50)$$

This is the interaction between two atoms, however, if there are many atoms in, for example, a crystal structure, one would want to include all interactions:

$$\hat{H} = - \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \quad (5.51)$$

this is the **Heisenberg model**.

However, for many materials $\langle r \rangle < d$ where d is the interatomic spacing, which implies that $S \approx 0$. Furthermore the large interatomic distance would imply that \hat{H}_{int} is negligible and hence $V \approx 0$: we cannot explain the magnetic properties of some solids due to the direct exchange (exchange between neighbouring atoms). In many materials it is necessary to consider some kind of indirect exchange interaction.

Superexchange

In a number of ionic solids there is an indirect exchange, where the electrons "communicate" non-locally: superexchange. Non-neighbouring ions are experiencing the exchange interaction, because the electrons delocalise to increase their kinetic energy. Let us consider a simple cubic crystal structure that has two atoms in the unit cell, say O and Mn. Mn has a single unpaired electron and O has a pair of electrons. Due to this superexchange is it beneficial for the next-to neighbouring Mn atoms to have anti-aligned spins, such that their electrons can delocalise without the Pauli exclusion principle "acting". The oxygen atom mediates the interaction between the unpaired electrons on the manganese atoms.

Indirect exchange in metals

In metals the conduction electrons mediate the exchange interaction between metallic ions. This is the RKKY interaction (also known as the itinerant exchange). We will show later that

$$J \sim r^{-3} \cos(2k_F r) \quad (5.52)$$

5.5 ORDER AND MAGNETIC STRUCTURES

Ferromagnetism

For a ferromagnet in a magnetic field the Hamiltonian is

$$\hat{H} = - \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + g\mu_B \sum_j \mathbf{S}_j \cdot \mathbf{B} \quad (5.53)$$

For magnetic ordering we would require that the nearest neighbour exchange constant is positive. The first term is very difficult to evaluate, however, for ferromagnetism we can use the Weiss model, which defines an effective molecular field at the i -th site:

$$\mathbf{B}_{\text{mf}} = - \frac{2}{g\mu_B} \sum_j J_{ij} \mathbf{S}_j \quad (5.54)$$

Which simplifies the Hamiltonian to

$$\hat{H} = g\mu_B \sum_i \mathbf{S}_i \cdot (\mathbf{B} + \mathbf{B}_{\text{mf}}) \quad (5.55)$$

The molecular field is due to the alignment of magnetic moments, and therefore we can assume that $\mathbf{B}_{\text{mf}} = \lambda \mathbf{M}$. The analysis in the book ignores L until now. But it is clear that S and L behave the same way in a magnetic field, therefore we can replace $\mathbf{S}_i \rightsquigarrow \mathbf{J}_i$. This brings us back to the Brillouin function:

$$M = M_s B_J(y), \quad y \equiv \frac{gJ\mu_B J(B + \lambda M)}{k_B T} \quad (5.56)$$

This can only be solved numerically. However, suppose that the magnetic energy is much smaller than the thermal energy: $\mu_B(B + \lambda M) \ll \frac{k_B T}{gJ}$ and hence we can use the first order approximation of the Brillouin function:

$$M \approx M_s \frac{gJ\mu_B(J+1)(B + \lambda M)}{3k_B T} \quad (5.57)$$

above the Curie-temperature, $T \geq T_C$, where

$$T_C = \frac{gJ\mu_B(J+1)\lambda M_s}{3k_B} \quad (5.58)$$

hence

$$M \approx \frac{T_C}{T} \left(\frac{B}{\lambda} + M \right) \rightsquigarrow M = \frac{T_C B}{\lambda(T - T_C)} \quad (5.59)$$

For small B it holds that $\chi = \frac{\mu_0 M}{B}$, which implies that

$$\chi \sim \frac{1}{T - T_C}, \quad T > T_C \quad (5.60)$$

which is the **Curie-Weiss** law.

Plotting the numerical solution to $\frac{M}{M_s}$ as a function of $\frac{T}{T_C}$ will give a graph that looks similar to the first quadrant of the unit circle (though not quite circular). This shows that the magnetisation will decrease as a function of $\frac{T}{T_C}$, and reach zero at $T = T_C$. The magnetisation is zero for $T \geq T_C$, which hints toward the fact that the magnetisation is an order parameter for magnetism.

Origin of the molecular field

The molecular field is what gives rise to spontaneous magnetisation, and hence ferromagnetism. Weiss could not obtain a large enough value of λ to describe nature, because he only considered dipole fields. It is, as we've seen, the exchange interaction that drives the spontaneous magnetisation. Using our description above we can show that

$$\lambda = \frac{2zJ}{ng^2\mu_B^2} \quad (5.61)$$

where z is the number of neighbouring atoms, which implies that

$$T_C = \frac{2zJ(J+1)}{3k_B} \quad (5.62)$$

Antiferromagnetism

The description of antiferromagnetism requires us to overlay two ferromagnetic lattices.

$$B_{\pm} = -|\lambda|M_{\mp} \quad (5.63)$$

Assuming the sub-lattices are identical we have that $|M_+| = |M_-| = M$, and hence

$$M = M_s B_J \left(\frac{g_J \mu_B J |\lambda| M}{k_B T} \right) \quad (5.64)$$

just like before. However, due to the opposite sign of λ , which now is negative, we must switch the sign of T_C everywhere: let us define the **Néel Temperature** as $T_N = -T_C$ for anti-ferromagnets:

$$\chi \sim \frac{1}{T + T_N} \quad (5.65)$$

This leads to a generalisation, where we instead use the *Weiss temperature*, θ :

$$\chi \sim \frac{1}{T - \theta}, \quad \theta \begin{cases} > 0 : & \text{ferromagnetism} \\ = 0 : & \text{paramagnetism} \\ < 0 : & \text{anti-ferromagnetism} \end{cases} \quad (5.66)$$

Ferrimagnetism

This is very similar to anti-ferromagnetism, however now the sublattices are not equivalent, and hence the atoms can have different magnetic moments. This means that the magnetisation can be non-zero, even though there are as many magnetic moments pointing upwards as there are moments pointing downwards.

Helical Order

In many rare earth metals the atoms lie in planes, which can lead to helical order, where the ferromagnetic planes are rotated an angle θ with respect to one another:

$$\cos \theta = -\frac{J_1}{4J_2} \quad (5.67)$$

where J_1 is the nearest neighbour coupling and J_2 is the next-to-nearest neighbour coupling

5.6 LANDAU THEORY

A macroscopic model that doesn't consider the microscopic information about the field, such as individual magnetic moments.

Landau theory is a mean-field theory that describes the Helmholtz free energy, F , as a polynomial expansion.

Let us first discuss the Helmholtz free energy and why a system will minimise it. We know from thermodynamics that

$$dS_{\text{universe}} = dS_{\text{system}} + dS_{\text{surroundings}} \geq 0 \quad (5.68)$$

For constant volume and particle number this gives us that

$$dS_{\text{universe}} = dS_{\text{system}} - T^{-1}dU_{\text{system}} = -T^{-1}dF_{\text{system}} \quad (5.69)$$

Thus for constant temperature, we maximise the universe's entropy by minimising the system's free energy.

In the case of ferromagnetism the order parameter is the magnetisation, M . The Helmholtz free energy is invariant under time-reversal, therefore we can conclude that the expansion only includes even terms of M . The simplest expansion that produces a *global* minimum that is non-trivial is

$$F = F_0 + aM^2 + bM^4 \quad (5.70)$$

where $a = a_0(T - T_C)$ and b : this ensures that F has a global minimum below the Curie-temperature, which will allow for magnetic ordering: non-zero magnetisation. The linear expansion of $a(T)$ naturally only holds near the Curie-temperature. Equation 5.70 is minimised for:

$$M \in \left\{ 0, \pm \sqrt{\frac{a_0(T_C - T)}{2b}} \right\} \quad (5.71)$$

In order to evaluate which of these minima are the global minimum we must calculate the free energy:

$$F(M = 0) = F_0, \quad F(\pm M_0) = F_0 - \frac{a_0^2}{4b}(T - T_C)^2 \quad (5.72)$$

thus the non-zero are equally favourable. The system has to spontaneously choose a magnetisation direction.

Using this theory we can also find an expression for the susceptibility of ferromagnets, by including another term:

$$F = F_0 + aM^2 + bM^4 - HM \quad (5.73)$$

which has different minima above and below the Curie-temperature:

$$\chi = \begin{cases} \frac{1}{2a_0(T - T_C)} & T > T_C \\ \frac{1}{4a_0(T_C - T)} & T < T_C \end{cases} \quad (5.74)$$

The entropy under constant volume and applied field is given by

$$S = -\frac{\partial F}{\partial T} = \begin{cases} S_0 & T > T_C \\ S_0 - \frac{a_0^2(T_C - T)}{2b} & T < T_C \end{cases} \quad (5.75)$$

This corresponds to $\mathbf{J} \cdot \mathbf{B}$ in our Hamiltonian: the free energy is minimised if the magnetisation points in the same direction as the applied field

and from here we can calculate the heat capacity:

$$c = T \frac{\partial S}{\partial T} = \begin{cases} c_0 \\ c_0 + \frac{a_0^2 T}{2b} \end{cases} \quad (5.76)$$

This is a second order phase transition, because there is a discontinuity in the second derivative of the Helmholtz free energy.

A first order transition can be described by including

$$F = F_0 + aM^2 + bM^4 + cM^6 \quad (5.77)$$

and we now require $c > 0$ and that b changes sign at some temperature. At this temperature there will be a first order transition.

5.7 MEAN-FIELD THEORY

As we discussed previously, we can simplify the Heisenberg model, by assuming that the spin-spin interaction can be described using a mean field, B_{mf} :

$$\hat{H} = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j \approx -\frac{2JzM}{g\mu_B n} \sum_i S_i^z \quad (5.78)$$

where we've let $\mathbf{B}_{mf} \parallel \mathbf{z}$ and z is the number of nearest neighbours. The mean-field is given by

$$B_{mf} = -\frac{2JzM}{ng^2\mu_B^2} \quad (5.79)$$

and hence we once again get a self-consistent equation:

$$M = \frac{g\mu_B n}{2} \tanh\left(\frac{\beta JzM}{ng\mu_B}\right) \quad (5.80)$$

Let us now solve this equation up to third order:

$$M = \frac{g\mu_B n}{2} \left(\frac{\beta JzM}{ng\mu_B} - \frac{1}{3} \left(\frac{\beta JzM}{ng\mu_B} \right)^3 \right) \quad (5.81)$$

$$\rightsquigarrow M = \frac{ng\mu_B T}{2T_C} \sqrt{\frac{3(T_C - T)}{T_C}} \quad (5.82)$$

This predicts $M \sim \sqrt{T_C - T}$ near the Curie-temperature, just as Landau theory did. Using the magnetisation we can now calculate the partition function, and from there other thermodynamic quantities. For instance:

$$\langle E \rangle = -\frac{k_B T M}{M_s} \tanh\left(\frac{T_C M}{T M_s}\right), \quad M_s = \frac{g\mu_B n}{2} \quad (5.83)$$

and

$$c = k_B \left(\frac{T_C M}{M_s} \right)^2 \operatorname{sech}^2\left(\frac{T_C M}{T M_s}\right) \quad (5.84)$$

An interesting side note is that using the mean-field Hamiltonian in the partition function gives us the free energy, whose minimum is exactly when Equation 5.80 holds.

Goldstone modes and the energy dispersion

Let us solve the Heisenberg Hamiltonian for a 1D chain, using the Fourier transformation. Let us consider a 1D chain of N moments, and furthermore let $|j\rangle = S_j^-|0\rangle$: we excited the ferromagnetic groundstate by lowering a spin at the j -th site. First we need that

$$S^-|SS\rangle = \sqrt{2S}|SS-1\rangle \quad (5.85)$$

$$S^+|SS-1\rangle = \sqrt{2S}|SS\rangle \quad (5.86)$$

and defining the Fourier transform of $|j\rangle$:

$$|q\rangle = \frac{1}{\sqrt{N}} \sum_j e^{iqr_j} |j\rangle \quad (5.87)$$

we get that

$$H|q\rangle = 2JS(-NS + 2(1 - \cos(qa)))|q\rangle \quad (5.88)$$

The long wave spin waves that emerge for small q are **Goldstone modes**:

Goldstone modes are long wavelength modes that appear when a continuous symmetry is broken spontaneously.

Our continuous symmetry is the rotational symmetry of M , which is broken spontaneously when the material "chooses" a magnetisation direction.

*Most likely because
 $\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1$ and only
1 has spin.*

The particle that carries the spin waves is the **magnon**, which is a boson. Therefore the relevant distribution function for magnons is the Bose-Einstein distribution function, from which we can calculate the total density of magnons:

$$n_{\text{magnon}} = \int_0^\infty \frac{g(\omega)d\omega}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1} \quad (5.89)$$

Converting this integral into a dimensionless integral:

$$n_{\text{magnon}} = \left(\frac{k_B T}{\hbar}\right)^{\frac{3}{2}} \int_0^\infty \frac{x^{\frac{1}{2}}}{e^x - 1} \sim T^{\frac{3}{2}} \quad (5.90)$$

Each magnon reduces total magnetisation by $g_s \mu_B$ (they are bosons remember), therefore:

$$\frac{M(0) - M(T)}{M(0)} \sim T^{\frac{3}{2}} \quad (5.91)$$

which is **Bloch's $T^{\frac{3}{2}}$ law**.

5.8 MAGNETISM IN METALS

If we ignore electron-electron interactions we can describe the electrons in metals using the **nearly free electron model**, which treats the lattice as a perturbation to the free electron Hamiltonian.

In 3D the density of states is:

$$g(k)dk = \left(\frac{L}{2\pi}\right)^3 8\pi k^2 dk = \frac{V k^2}{\pi^2} dk \quad (5.92)$$

where the additional factor two comes from spin degeneracy. From here we can calculate the dependence of N on the fermi energy at $T = 0$:

$$N = \int_0^{k_F} dk g(k) = \frac{V k_F^3}{3\pi^2} \rightsquigarrow k_F^3 = 3\pi^2 n \quad (5.93)$$

Naturally

$$E_F = \frac{\hbar^2 k_F^2}{2m} \rightsquigarrow g(E_F) = \frac{3}{2} \frac{n}{E_F} = \frac{3}{2} \frac{k_F^3}{3\pi^2 E_F} \quad (5.94)$$

For $T > 0$ the implicit Heavyside function used in the integral for N becomes a Fermi-Dirac distribution function:

$$f_{\text{FD}}(E, \mu) = \frac{1}{e^{\frac{(E-\mu)}{k_B T}} + 1} \quad (5.95)$$

where $\lim_{T \rightarrow 0} E_F(T) = 0$. When $f_{\text{FD}}(E, \mu) \approx \Theta(E - \mu)$, which is nearly always the case for metals that are below their melting temperature, it is said that the electrons are in the *degenerate limit*. In the non-degenerate limit, which corresponds to $(E - \mu) \gg k_B T$ the Fermi-Dirac distribution becomes asymptotic to the Boltzmann distribution $e^{-\frac{(E-\mu)}{k_B T}}$. The Fermi energy, E_F is equal to the chemical potential at $T = 0$, but for non-zero temperatures we must use the Sommerfeld approximation to find that:

$$\mu(T) = E_F \left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 + O \left(\left(\frac{k_B T}{E_F} \right)^4 \right) \right) \quad (5.96)$$

Pauli Paramagnetism

Electrons are negatively charged and therefore their magnetic moments point opposite their spin. Which means that if the magnetic moment aligns with the magnetic field, then the spin anti-aligns. Let the magnetic field point "downwards" so that it is favourable for spins to point upwards. Therefore the application of a magnetic field will move some spins from the down state to the up state:

$$n_{\uparrow} = \frac{1}{2} \int_0^{\infty} g(E + \mu_B) f_{\text{FD}}(E) dE \quad (5.97)$$

$$n_{\downarrow} = \frac{1}{2} \int_0^{\infty} g(E - \mu_B) f_{\text{FD}}(E) dE \quad (5.98)$$

The magnetisation is

$$M = \mu_B (n_{\uparrow} - n_{\downarrow}) \approx \mu_B^2 B \int_0^{\infty} \frac{dg}{dE} f_{\text{FD}}(E) dE \quad (5.99)$$

where we have used that B is small and hence the difference can be approximated as the derivative of g . Now, using integration by parts and the fact that the boundary terms is equal to zero we get

$$M \approx -\mu_B^2 B \int_0^{\infty} \frac{df}{dE} g(E) dE \quad (5.100)$$

In the degenerate limit $f' = \delta(E - E_F)$:

$$\chi = \frac{\mu_0 M}{B} = \mu_0 \mu_B^2 g(E_F) \quad (5.101)$$

whereas in the non-degenerate limit:

$$-\frac{df}{dE} = \frac{f}{k_B T} \quad (5.102)$$

and hence

$$\chi = \frac{n \mu_0 \mu_B^2}{k_B T} \quad (5.103)$$

following the Curie-law.

Spontaneously split bands

Let us consider a ferromagnetic material, where the molecular field is given by λM . Blundell argues that this molecular field will magnetise the material further, which may lead to spontaneous ferromagnetism. Blundell continues to try and establish whether (or how) a ferromagnetic system can save energy by being in the ferromagnetic state.

The potential energy is

$$\Delta V = - \int_0^M \mu_0 \lambda M' dM' = -\frac{1}{2} \mu_0 \lambda M^2 = -\frac{1}{2} \mu_0 \mu_B^2 \lambda (n_\uparrow - n_\downarrow)^2 \quad (5.104)$$

Let us write that as

$$\Delta V = -\frac{1}{2} U g^2(E_F) \delta E^2 \quad (5.105)$$

where U is a measure of the Coulomb interaction and δE is the energy gained by spin flipping. The "kinetic" energy is

$$\Delta T = \frac{1}{2} g(E_F) \delta E^2 \quad (5.106)$$

Thus the total change in energy is

$$\Delta E = \frac{1}{2} g(E_F) \delta E^2 (1 - U g(E_F)) \quad (5.107)$$

therefore spontaneous ferromagnetism is beneficial if $U g(E_F) > 1$, which is known as the **Stoner criterion**. From this energy shift we can calculate the susceptibility:

$$\chi = \frac{\chi_P}{1 - U g(E_F)} \quad (5.108)$$

For $1 - U g(E_F) < 1$ we have a so-called *Stoner Enhancement*.

LANDAU LEVELS

Let us consider a free electron in a magnetic field. Using the symmetric gauge, the vector potential for a field point along the z axis is:

$$\mathbf{A} = Bxy \quad (5.109)$$

thus the Schrödinger Equation is:

$$\frac{1}{2m_e} \left(\hbar^2 \frac{\partial^2}{\partial x^2} + \left(\hbar^2 \frac{\partial^2}{\partial y^2} + eBx \right) + \hbar^2 \frac{\partial^2}{\partial z^2} \right) \psi = E\psi \quad (5.110)$$

We still have that $[p_y, H] = 0$ and $[p_z, H] = 0$, therefore we can write our wavefunction as

$$\psi = u(x)e^{i(k_y y + k_z z)} \quad (5.111)$$

which means

$$\hat{H} = \frac{p_x^2}{2m_e} + \frac{1}{2}m\omega_c^2(x - x_0)^2 + \frac{\hbar^2 k_z^2}{2m_e} \quad (5.112)$$

where we've defined

$$\omega_c = \frac{eB}{m_e}, \quad x_0 = -\frac{\hbar k_y}{eB} \quad (5.113)$$

This is the dispersion of a free electron and a harmonic oscillator:

$$E = \left(\ell + \frac{1}{2} \right) \hbar \omega_c + \frac{\hbar^2 k_z^2}{2m_e}, \quad \ell \in \mathbb{N}^0 \quad (5.114)$$

It is not visible in this formula, but the system has cylindrical symmetry: the states form Landau tubes, parallel to the magnetic field in reciprocal space.

Landau Diamagnetism

The density of state for Landau levels is

$$g(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \frac{\hbar \omega_c}{2} \sum_{n=0}^{n_{\max}} \left(E - \left(n + \frac{1}{2} \right) \hbar \omega_c \right)^{\frac{1}{2}} \quad (5.115)$$

Using this in $E = \int \varepsilon g(\varepsilon) d\varepsilon$ we get that

$$E = \frac{3NE_F}{5} - \frac{V}{24\pi^2} e^2 B^2 \frac{k_F}{2m_e} \quad (5.116)$$

which in turn gives us a *negative* contribution to the susceptibility:

$$\chi_{\text{Landau}} = \frac{\mu_0}{V} \frac{\partial^2 E}{\partial B^2} = -\frac{\mu_0 e^2 k_F}{3m_e} \frac{1}{4\pi^2} = -\frac{\mu_0 \mu_B^2}{3} g(E_F) \quad (5.117)$$

precisely one third of the value of the Pauli susceptibility.

Momentum-dependent susceptibility

Let us apply a sinusoidal magnetic field to an itinerant gas:

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_{\mathbf{q}} \cos(\mathbf{q} \cdot \mathbf{r}) \quad (5.118)$$

Leading to a perturbation:

$$\hat{H}' = \pm \frac{g\mu_0 \mu_B}{2} |\mathbf{H}_{\mathbf{q}}| \cos(\mathbf{q} \cdot \mathbf{r}) \quad (5.119)$$

The first order correction to the wave function is

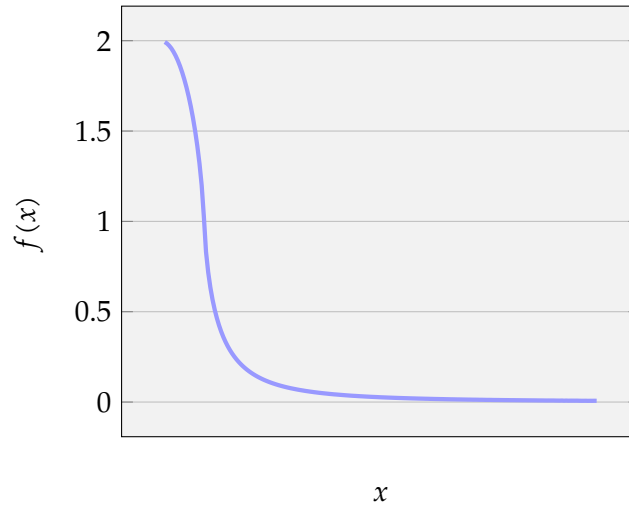
$$\frac{1}{\sqrt{V}} \left(e^{i\mathbf{k}\cdot\mathbf{r}} \pm \frac{g\mu_0\mu_B H_{\mathbf{q}}}{4} \right) \left(\frac{e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}}}{E_{\mathbf{k}+\mathbf{q}} - E_{\mathbf{k}}} + \frac{e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{r}}}{E_{\mathbf{k}-\mathbf{q}} - E_{\mathbf{k}}} \right) |\pm\rangle \quad (5.120)$$

this can be used to get an expression for the magnetisation – the norm squared of the spin up and spin down electrons is proportional to the density. Finally we integrate over \mathbf{k} -space, which removes the k -dependence, and we end up with a susceptibility:

$$\chi_{\mathbf{q}} = \chi_P f\left(\frac{q}{2k_F}\right) \quad (5.121)$$

where

$$f(x) = \frac{1}{2} \left(1 + \frac{1-x^2}{2x} \log \left| \frac{1+x}{1-x} \right| \right) \quad (5.122)$$

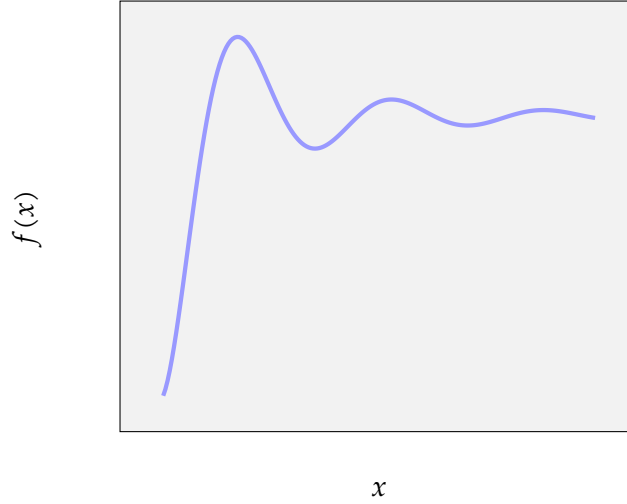


RKKY-INTERACTION

We have the susceptibility's behaviour in reciprocal space, so let us calculate the inverse Fourier transform:

$$\chi(\mathbf{r}) = \mathcal{F}^{-1} \{ \chi_{\mathbf{q}} \} = \frac{1}{(2\pi)^3} \int d^3\mathbf{q} \chi_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \quad (5.123)$$

$$= \frac{2k_F^3 \chi_P}{\pi} F(2k_F r) \quad (5.124)$$



Spin density waves

Just like for the electron gas treated with q -independent excitations, the q -dependent susceptibility experiences a Stoner enhancement:

$$\chi_{\mathbf{q}} = \frac{\chi_{\mathbf{q}}^0}{1 - \alpha \chi_{\mathbf{q}}^0} \quad (5.125)$$

where $\chi_{\mathbf{q}}^0$ is the expression we found previously. It may happen that $\chi_{\mathbf{q}}^0$ has a maximum at a non-zero value of \mathbf{q} , in which case it is possible for spontaneous magnetisation to appear in the sample and we would expect spiral spin-density waves in the wave vector, \mathbf{q} . The derivation of the q dependence of $\chi_{\mathbf{q}}$ can be repeated in 1D and 2D: the 1D susceptibility diverges at $q = 2k_F$, which is the **Kohn anomaly**, which shows that one dimensional electron gasses are unstable to the formation of spin density waves with wave vector $2k_F$. This is associated with the fact that the modulation opens up a gap at the fermi surface, which allows states with $q = 2k_F$ to lower their energy, which is an opportunity to drive the formation of a density wave, which causes it to diverge.

Kondo Effect

Below some temperature, the **Kondo temperature**, T_K , electrons screen the magnetic moments of nuclei/ions. This shields any magnetic properties of the material, and also creates quasi-bound states. These bound states increase the scattering area, which means that the resistivity has a term $-|J| \ln T$ which diverges at low temperatures

Hubbard Model

In some materials the electron-nucleus binding energy is so strong, that electrons cannot be treated as (nearly) free electrons. In this case we must use the Hubbard model, where we treat the electrons as excitations of each site, which hopping amplitudes, such that the electrons can move through the material if the Pauli-exclusion principle allows this.

$$\hat{H} = -t \sum_{\sigma, \langle ij \rangle} (c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma}) + U \sum_i (n_{i\uparrow} + n_{i\downarrow} - 1)^2 \quad (5.126)$$

$$- \mu \sum_i (n_{i\uparrow} + n_{i\downarrow}) \quad (5.127)$$

The first term is hopping between sights, and the second term is the Coulomb energy, which will try to move the electrons away from each other. Setting $\mu = 0$ and treating the first term as a perturbation of the second term, you can show that the effective Hamiltonian (first order) can be written as

$$H_{eff} = \frac{2t^2}{U} \sum_{\langle ij \rangle} \left(\mathbf{S}_i \cdot \mathbf{S}_j + \frac{1}{8} \right) \quad (5.128)$$

5.9 SUPERCONDUCTIVITY

For conductive (linear) materials

$$\mathbf{j} = \sigma \mathbf{E}, \quad \Leftrightarrow \quad \mathbf{E} = \rho \mathbf{j} \quad (5.129)$$

Generally $\rho(T)$ is an increasing function, because phononic excitations make the crystal lattice "less uniform" which increases scattering rates. For superconductors there exists a critical temperature, T_C , where

$$\rho(T) = 0, \quad T < T_C \quad (5.130)$$

However, this is not the *defining* characteristic, because there exists materials that have $\rho = 0$ and $\sigma = 0$. The defining characteristic of superconductors is the **Meissner effect**, which will be described in further detail shortly. The super-currents in the surface of the super conductor shield any exterior magnetic field, such that $\mathbf{B}_{int} = 0$. This implies that $\chi = -1$. However, there exists a critical field, \mathbf{H}_C where the superconductor begins to behave like a normal conductor again.

Choosing the London Gauge, $\nabla \cdot \mathbf{A} = 0$, we can write

$$\mathbf{j} = -\frac{n_e e^2}{m} \mathbf{A} \quad \rightsquigarrow \quad \nabla^2 \mathbf{B} = \frac{1}{\lambda_L^2} \mathbf{B}, \quad \lambda_L \equiv \sqrt{\frac{m}{\mu_0 n_s e^2}} \quad (5.131)$$

λ_L is the surface depth of the superconductor. The current carriers in superconductors are not electrons, but Cooper pairs, that are bosons.

Thermodynamics of Superconductors

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) = 0 \quad \rightsquigarrow \quad \chi = -1 \quad (5.132)$$

However, this only holds up to the critical temperature. Type-I superconductors have a discontinuity in χ at $H = H_C$, where the susceptibility abruptly falls to zero. Type-II superconductors have two critical temperatures; H_{C1} where χ transitions to zero continuously. Above H_{C1} , M decays in the "mixed region" where there are part that are superconducting and parts that are in the normal state.

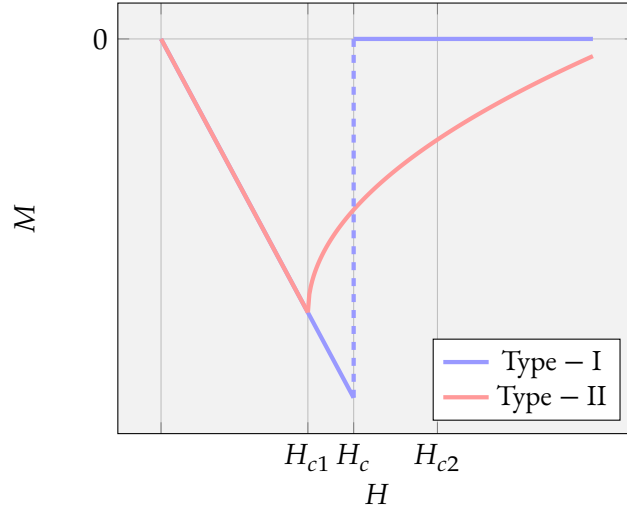


Figure 5.1: Schematic drawing of the magnetisation as a function of applied field for type-I and type-II superconductors. In the region $H_{c1} < H < H_{c2}$ vortices can exist in the type-II super conductors.

For an *isothermal* type-I superconductor:

$$G = -\mu_0 \int_0^H M(H') dH' + G_0 \quad (5.133)$$

thus

$$G_s(T, H) = G_s(T, 0) + \frac{\mu_0 H^2}{2} \quad (5.134)$$

The phase transition occurs at $H = H_c$, therefore at this applied field we have that

$$G_n(T) - G_s(T, H_c) = 0 \quad \rightsquigarrow \quad G_n(T) - G_s(T, 0) = -\frac{\mu_0 H_c^2}{2} \quad (5.135)$$

which just reflects that it is more beneficial to be in the superconducting phase below the critical field.

The phase transition is a second order transition, because the discontinuity is in the specific heat:

$$c_s - c_n|_{T=T_c} = 4\mu_0 \frac{H_c^2}{T_c} \quad (5.136)$$

Due to this discontinuity it is sometimes easier to find the phase transition using c , than find where $\rho \rightarrow 0$.

Ginzburg-Landau Theory

The order parameter for superconductivity is $\psi(r) = |\psi(r)|e^{i\phi(r)}$, which can be thought of as a kind of macroscopic wave function. The free energy density is then:

$$f_s = f_n + \alpha(T)|\psi(r)|^2 + \frac{1}{2}\beta|\psi(r)|^4 \quad (5.137)$$

$$+ \frac{1}{2m^*} |(-i\hbar\nabla + e^*\mathbf{A})\psi|^2 + \frac{B^2}{2\mu_0} \quad (5.138)$$

This free energy is symmetric under Gauge Transformations:

$$\{\psi(r), \mathbf{A}(r)\} \rightsquigarrow \left\{ \psi(r)e^{i\theta(r)}, \mathbf{A} + \frac{\hbar}{e^*}\nabla\theta(r) \right\} \quad (5.139)$$

This symmetry is broken spontaneous below the critical temperature: the superconductor "picks" a specific phase.

$|\psi|$ is proportional to the density of Cooper-pairs.

Let us set $B = 0$ and minimise the free energy and only consider that homogeneous case:

$$\delta f_s = (\alpha(T)\psi(r) + \beta|\psi(r)|^2\psi) \delta\psi^* + \dots \quad (5.140)$$

which gives us

$$|\psi|^2 = \frac{-a_0(T - T_c)}{\beta} \quad (5.141)$$

Using this and the result for the Gibbs energy difference above, we can show that

$$H_c(T) = \frac{a_0}{\sqrt{\mu_0 b}}(T_c - T) \quad (5.142)$$

near T_c . For the non-homogenous case, we have the ∇ term, which just adds on to the equation, making it look similar to a Schrödinger equation, except that it is non-linear:

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi(r) + a(T)\psi(r) + b|\psi(r)|^2\psi(r) = 0 \quad (5.143)$$

This is the **first Ginzburg-Landau equation**. For instance, this tells us about the attenuation of the order parameter close to the interface to (let us say) a vacuum:

$$\psi(x) = \psi_0 \tanh\left(\frac{x}{\sqrt{2}\zeta(T)}\right), \quad \text{for } x \geq 0 \quad (5.144)$$

where the constant, ζ , is the *coherence length* or *healing length*.

$$\zeta(T) = \sqrt{\frac{\hbar^2}{2m^*|a(T)|}} \quad (5.145)$$

For $B \neq 0$ we can just replace $\hat{\mathbf{p}}$ with $\hat{\mathbf{p}} + e^*\mathbf{A}$. However, we can additionally require that the free energy is minimised with respect to A , which in case we get the **second Ginzburg-Landau equation**:

$$\mathbf{j}_s = \frac{\hbar e^*}{2m^*i} (\psi\nabla\psi^* - \psi^*\nabla\psi) - \frac{(e^*)^2}{m^*}|\psi|^2\mathbf{A} \quad (5.146)$$

this looks exactly like the quantum mechanical current, however describes a macroscopic quantity. For any shape that is topological to a ball, if we assume that $|\psi| = \text{const.}$ and that there is a steady state, in which case $\nabla \cdot \mathbf{j}_s = 0$, we get that

$$\mathbf{j}_s = -\frac{(e^*)^2}{m^*}|\psi|^2\mathbf{A} \quad (5.147)$$

Another interesting artefact of superconductivity is **flux-quantisation**. Consider a doughnut geometry and again assume the order parameter only varies in its phase, then it can be shown that

$$F_s = F_s^0 + \Gamma(\Phi - n\Phi_0)^2 + \Omega\Phi^2, \quad n \in \mathbb{Z}, \quad \Phi_0 = \frac{h}{2e} \quad (5.148)$$

which is minimised for $\Phi = n\Phi_0$, $n \in \mathbb{Z}$: exactly when the flux is an integer multiple of the quantum flux. Another way of approaching the description of flux-quantisation is through the London Equation:

$$\oint \mathbf{j}_s \cdot d\ell = \frac{e^*}{m^*} |\psi|^2 \oint (h\nabla\psi - e^* \mathbf{A}) \cdot d\ell \quad (5.149)$$

Inside the doughnut $\mathbf{j}_s = \mathbf{0}$, so

$$\oint \mathbf{A} \cdot d\ell = \frac{h}{4\pi e} \oint \nabla\phi \cdot d\ell = n \left(\frac{h}{2e} \right) = n\Phi_0, \quad n \in \mathbb{Z} \quad (5.150)$$

Assuming that ϕ is single valued.

Type-I and Type-II superconductivity

The two scale lengths

$$\lambda_L = \sqrt{\frac{m^* \beta}{\mu_0 e^* a_0 (T - T_C)}} \quad (5.151)$$

$$\xi = \sqrt{\frac{\hbar^2}{2ma_0(T_c - T)}} \quad (5.152)$$

are sufficient to distinguish the two types of superconductors:

$$\kappa \equiv \frac{\lambda_L}{\xi} \quad \begin{cases} < \frac{1}{\sqrt{2}} & \text{type - I} \\ > \frac{1}{\sqrt{2}} & \text{type - II} \end{cases} \quad (5.153)$$

Little-Parks Experiment

The experiment studies a thin superconducting ring, where the thickness, $d \approx \lambda_L, \xi$. The critical temperature oscillates:

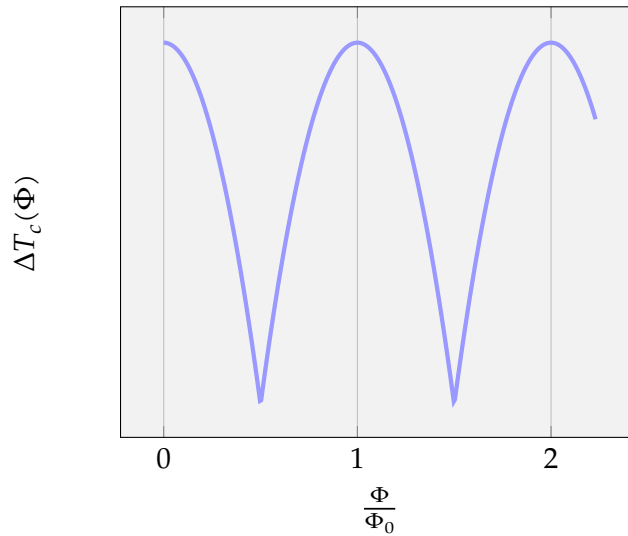


Figure 5.2: Change in critical temperature as a function of magnetic flux in the Little-Parks experiment. The amplitude of the oscillation is approximately 1mK.

This result can be explained with Ginzburg-Landau theory. Consider the current in the superconductor:

$$\mathbf{j}_s = \frac{e^*}{m^*} |\psi|^2 (\hbar \nabla \theta - e^* \mathbf{A}) = en_s \mathbf{v}_s \quad (5.154)$$

The movement of cooper pairs corresponds to a kinetic energy, which should be included in the free energy:

$$f_s = f_n + \alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + |\psi|^2 2m^* \|\mathbf{v}_s\|^2 + \frac{\|\mathbf{B}\|^2}{2\mu_0} \quad (5.155)$$

At equilibrium ($\delta f_s = 0$) and hence

$$|\psi|^2 = -\frac{\alpha + 2m^* \|\mathbf{v}_s\|^2}{\beta} = |\psi_0|^2 \left(1 - \frac{2m^* \|\mathbf{v}_s\|^2}{|\alpha|} \right) \quad (5.156)$$

for $T < T_c$. By plugging this into the old expression for the free energy and comparing to the term above, we see that

$$v_s = \frac{\hbar}{2m^* R} \left(n - \frac{\Phi}{\Phi_0} \right) \quad (5.157)$$

Additionally we have that

$$\mathbf{j}_s = 2e^* |\psi_0|^2 \left(1 - \frac{2m^* \|\mathbf{v}_s\|^2}{|\alpha|} \right) \mathbf{v}_s \quad (5.158)$$

at equilibrium

$$-|\alpha| = a_0 (T - T_c) = -\frac{\hbar^2}{2m^* R^2} \left(n - \frac{\Phi}{\Phi_0} \right)^2 \quad (5.159)$$

Interface & Surface Energies

The normal and superconducting phases can mix, or not mix, depending on whether the surface is positive or negative. The positive case means that the surface keeps to itself, separate from the superconducting phase, whereas negative surface energy implies that the surface states are mixed into the superconductor. The latter is type-II superconductivity.

The Gibbs free energy can be obtained with a Legendre transformation:

$$G = F - HB \quad (5.160)$$

according to Ginzburg-Landau theory, this would imply that

$$g_n(T, H) = f_n(T, 0) + \frac{1}{2\mu_0} B^2 - HB \quad (5.161)$$

$$= f_n(T, 0) - \frac{1}{2} \mu_0 H^2 \quad (5.162)$$

because the order parameter vanishes in the normal state. For superconducting state, however, it is:

$$g_s(T, H) = f_n(T, 0) + \alpha(T) |\psi|^2 + \frac{\beta}{2} |\psi|^4 \quad (5.163)$$

$$+ \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \nabla + e^* \mathbf{A} \right) \psi \right|^2 + \frac{1}{2\mu_0} B^2 - HB \quad (5.164)$$

This implies that the "surface energy", σ_{ns} , which is the difference between G_n and G_s is at $H = H_c$:

$$\sigma_{ns} = \int_{-\infty}^{\infty} \Delta G dx \quad (5.165)$$

$$= \int_{-\infty}^{\infty} \alpha(T) |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \nabla + e^* \mathbf{A} \right) \psi \right|^2 \quad (5.166)$$

$$+ \frac{(B - \mu_0 H_c)^2}{2\mu_0} dx \quad (5.167)$$

However it turns out that there are a few terms that integrate to zero, which simplifies to

$$\sigma_{ns} = \int_{-\infty}^{\infty} \left(-\frac{1}{2} \beta |\psi|^4 + \frac{(B - \mu_0 H_c)^2}{2\mu_0} \right) dx \quad (5.168)$$

If the first term dominates, the surface energy is negative, which implies that we have a type-II superconductor and it is energetically beneficial to disperse surfaces within the sample. Note that this is at $H = H_c$. At $H < H_c$ we'd expect positive surface energies in type-I and in type-II we'd expect positive surface energies for $H < H_{c1}$.

Type-II Superconductors

For strong H $|\psi|^2 \ll 1$, because the strong magnetic field is in the process of destroying the superconductivity. This means we can linearise the Ginzburg Landau equation:

$$\frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla + e^* \mathbf{A} \right)^2 \psi = -\alpha \psi \quad (5.169)$$

For $B \parallel z$ and using the Landau gauge, in which case $\mathbf{A} = Bx\mathbf{y}$ and this becomes the equation for a harmonic oscillator. Therefore we conclude that

$$a_0(T - T_c) = \hbar \omega_c \left(n + \frac{1}{2} \right) + \frac{\hbar^2 k_z^2}{2m^*}, \quad \omega_c = \frac{e^* B}{m^*} \quad (5.170)$$

This implies that

$$B = \frac{-\alpha m^*}{e^* \hbar \left(n + \frac{1}{2} \right)} \quad (5.171)$$

which is maximised for $n = 0$: the superconductor can't support stronger magnetic field and it goes into the normal state. Thus

$$H_{c2} = \frac{B}{\mu_0} = -\frac{\alpha m^*}{\hbar e \mu_0} = \frac{1}{\mu_0} \frac{\Phi_0}{2\pi \xi^2}, \quad \xi^2 = \frac{\hbar^2}{2ma_0(T_c - T)} \quad (5.172)$$

The first critical field is a bit more difficult to calculate. Let us calculate the energy of a single vortex. The current, just like the flux, is quantised, therefore we expect that the momentum is quantised such that

$$\oint \mathbf{p} \cdot d\mathbf{l} = nh \quad \rightsquigarrow v_s = \frac{n\hbar}{m^* r} \quad (5.173)$$

Now consider the kinetic energy in the region where the supercurrent is correcting for the rogue field, i.e. $\xi < r < \lambda_L$ giving us

$$E_{\text{vortex}} = n_s^* \int_{\xi}^{\lambda_L} \frac{\pi n^2 \hbar^2}{m^* r} dr = \frac{\pi n_s^* \hbar^2}{m^*} \ln \kappa \quad (5.174)$$

In that order, because this is a type-II superconductor.

The first vortex appear where

$$H_{c1} \Phi_0 = E_{\text{vortex}} \quad (5.175)$$

and hence

$$H_{c1} = \frac{\Phi_0^2}{4\pi\mu_0\lambda_L^2} \ln \kappa \quad (5.176)$$

Josephson junction:

Let us consider two superconductors separated by an insulator of thickness, d .

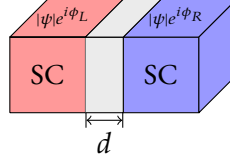


Figure 5.3: The grey area is an insulator situated between two superconductors with order parameters $|\psi|e^{i\phi_L}$ and $|\psi|e^{-i\phi_R}$.

The dimensionless (first) Ginzburg-Landau equation is

$$\xi^2 f'' + f - f^3 = 0 \quad (5.177)$$

but let us include a factor d^{-2} :

$$\left(\frac{\xi}{d}\right)^2 f'' + d^{-2}f - d^{-2}f^3 = 0 \quad (5.178)$$

for $d \ll \xi$ the first term dominates, giving us a Laplace equation, the solution to which can be approximated to

$$f(x) \approx \left(\frac{1}{2} - \frac{x}{d}\right) e^{i\phi_L} + \left(\frac{1}{2} + \frac{x}{d}\right) e^{i\phi_R} \quad (5.179)$$

as this is a solution to the Laplace equation. Now, inserting this expression into the second Ginzburg-Landau equation for $B = 0$ gives us

$$j = \frac{2e\hbar|\psi_0|^2}{m^*\alpha} \sin(\phi_R - \phi_L) \quad (5.180)$$

This implies that if we have two superconductors separated by an insulating layer and these superconductors have order parameters that only differ by a phase factor, then current will pass through the insulating layer, *even* when the voltage difference is zero.

This effect is closer described using Feynman's method, where we use that it is the chemical potential that sets the time-evolution of ψ :

*Is it a kind of
time-dependent
Ginzburg-Landau theory?*

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \mu \psi \quad (5.181)$$

However, now it is the case that the left and right order parameters are coupled. Let us couple them phenomenologically, using a first order approximation:

$$-\frac{\hbar}{i} \frac{d}{dt} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \mu_L & T \\ T & \mu_R \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad (5.182)$$

We get (by setting in and looking at the real and imaginary parts separately)

$$\frac{\partial|\psi_L|}{\partial t} = \frac{1}{\hbar}T|\psi_R|\sin(\phi_R - \phi_L) \quad (5.183)$$

This tells us about the current from the one side to the other:

$$j = -e^* \frac{\partial|\psi_L|^2}{\partial t} = -2e^*|\psi_L|\dot{|\psi_L|} \quad (5.184)$$

$$= -\frac{2e^*T}{\hbar}|\psi_L||\psi_R|\sin(\phi_L - \phi_R) \quad (5.185)$$

The time evolution can also be calculated from this "Schrödinger" equation:

$$\frac{\partial}{\partial t}(\phi_R - \phi_L) = \frac{1}{\hbar}(\mu_L - \mu_R) \quad (5.186)$$

which is constant when $\Delta V = 0$.

Non-zero voltage difference

In the case that $\Delta V \neq 0$ we instead get

$$\frac{\partial}{\partial t}(\phi_R - \phi_L) = \frac{2e\Delta V}{\hbar} \rightsquigarrow \Delta\phi = \Delta\phi_0 + \frac{2e\Delta V}{\hbar}t \quad (5.187)$$

which in turn implies that

$$j_s = j_0 \sin\left(\Delta\phi_0 + \frac{2e\Delta V}{\hbar}t\right) \quad (5.188)$$

This is a very high frequency:

$$\frac{2e\Delta V}{\hbar} = \frac{\Delta V}{0.33\text{mV}}\text{THz} \quad (5.189)$$

Josephson function in magnetic field

Consider a setup as before, but we now apply a magnetic field, parallel to the z-axis. The magnetic field attenuates exponentially within the superconductors, but remains constant in the insulating region. Suppose the interface is the yz -plane. Let us $\mathbf{A} = A(x)\mathbf{y}$.

Far inside the superconductor the current is zero, hence

$$\nabla\phi = -\frac{e^*}{\hbar}\mathbf{A} \quad (5.190)$$

This implies that the phase difference between two point within the same superconductor (still far inside) is given by:

$$\phi(L_2) - \phi(L_1) = \int_{L_1}^{L_2} \nabla\phi \cdot d\ell = -\frac{e^*}{\hbar} \int_{L_1}^{L_2} \mathbf{A} \cdot d\ell \quad (5.191)$$

And similarly far inside the right superconductor. Suppose line connecting L_i and R_i is parallel to the x -axis, and hence $\mathbf{A} \cdot d\ell = 0$. This implies we can add this quantity to the integral, without cost, hence:

$$\frac{e^*}{\hbar} \oint \mathbf{A} \cdot d\ell = 2\pi \frac{\Phi}{\Phi_0} \quad (5.192)$$

Let L_1 and R_1 be at $y = 0$ and L_2 and R_2 be at $y = y$, hence

$$\Delta\phi(y) - \Delta\phi(0) = 2\pi \frac{\Phi(y)}{\Phi_0}, \quad \Phi(y) = B(d + \lambda_L + \lambda_R)y \quad (5.193)$$

This means the current becomes y -dependent:

$$J_x(y) = J_0 \sin\left(\Delta\phi_0 + 2\pi \frac{\Phi(y)}{\Phi_0}\right) \quad (5.194)$$

Integrating over y :

$$I_f = I_y \sin(\delta\phi_0) \frac{\sin\left(\pi \frac{\phi}{\phi_0}\right)}{\pi \frac{\Phi}{\Phi_0}} \quad (5.195)$$

SQUIDS

Creating a loop that links two superconductors at two points, with two insulators creates a SQUID. This device is very sensitive to magnetic flux due to the Fraunhofer diffraction pattern. Thus this device measures flux in units of the flux-quantum!

LANDAU LEVELS: REVISITED

As described by Hamiltonian mechanics (and quantum mechanics), the application of a magnetic field to an electronic gas, corresponds to the transformation:

$$-i\hbar\nabla \rightarrow -i\hbar\nabla + e\mathbf{A} \quad (5.196)$$

Let us consider free electrons in a box, and apply a magnetic field $\mathbf{B} \parallel \mathbf{z}$. This can be solved in both the Landau and the symmetric gauge. We focussed on the Landau gauge:

$$\mathbf{A} = Bx\mathbf{y} \quad (5.197)$$

This, as before leads to a harmonic oscillator:

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega_c^2(x - x_0^2) + \frac{\hbar^2 k_z^2}{2m} \quad (5.198)$$

and an energy of

$$E = \left(\ell + \frac{1}{2}\right)\hbar\omega + \frac{\hbar^2 k_z^2}{2m} \quad (5.199)$$

k_y only decides the equilibrium position and hence does not influence the energy: degenerate in k_y . However, x_0 must be inside the sample:

$$0 < x_0 = \frac{\hbar k_y}{eB} < L_x \quad (5.200)$$

but the free electron gas has

$$\frac{2\pi n}{L_y} = k_y < L_x \quad (5.201)$$

and hence

$$n \in N, \quad N = \left\{ n \in \mathbb{N} : n < \frac{L_x L_y e B}{h} \right\} \quad (5.202)$$

We don't include $n = 0$, because that would mean that $x_0 \in \partial V$. Note that $L_x L_y B = \Phi$ and $\frac{eB}{h} = \Phi_0$ so

$$N_P = \frac{\Phi}{\Phi_0} \quad (5.203)$$

where $N_P = \left\lfloor \frac{L_x L_y e B}{h} \right\rfloor$ is the degeneracy. Thus adding one state to the degenerate space corresponds to adding a flux quantum to the flux.

Density of states

The n -th Landau subband has N_n states, which is N_P multiplied by the number of k_z values. The number of k_y values between k_y and $k_y + dk_y$ is $\frac{L_z}{2\pi} dk_z$, and therefore

$$dN_n = L_x L_y L_z \frac{2eB}{(2\pi)^2 \hbar} dk_z = \frac{2VeB}{(2\pi)^2 \hbar} dk_z \quad (5.204)$$

where the extra factor two is due to spin degeneracy. Substituting for dk_z :

$$dN_n = \frac{2VeB}{(2\pi)^2 \hbar} \frac{\sqrt{2m}}{\hbar} \left(E - \left(n + \frac{1}{2} \right) \hbar \omega_c \right)^{-\frac{1}{2}} dE \quad (5.205)$$

Now the additional factor two is because k can be negative, but E is strictly positive. So

$$g_n(E) = \frac{1}{V} \frac{dN_n}{dE} = \frac{\sqrt{2m}}{2\pi^2 \hbar^2} eB \left(E - \left(n + \frac{1}{2} \right) \hbar \omega_c \right)^{-\frac{1}{2}} \quad (5.206)$$

Summing

$$g(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \frac{\hbar \omega_c}{2} \sum_{n=0}^{n_{\max}} \left(E - \left(n + \frac{1}{2} \right) \hbar \omega_c \right)^{-\frac{1}{2}} \quad (5.207)$$

In reciprocal space the states of a free electron form a filled circle. When we apply a magnetic field it changes in to rings, with $\frac{\Phi}{\Phi_0}$ states per ring.

The magnetic field slices the fermi surface

5.10 INTEGER QUANTUM HALL EFFECT

Classically we had that

$$E_i = \rho_{ij} j_j \quad (5.208)$$

where the resistivity tensor is

$$\rho_{ij} = \rho_0 \delta_{ij} + \varepsilon_{ijk} H_k, \quad \rho_0 = \frac{m}{ne^2 \tau}, \quad H_k = \frac{B_k}{ne} \quad (5.209)$$

however, in the quantum limit (so low energy limit), we measure a quantisation of these quantities, that follow:

$$\rho_{xy} = \frac{h}{\xi e^2}, \quad \xi \in \mathbb{N} \quad (5.210)$$

and ρ_{xx} has a narrow spike every time ρ_{xy} transitions from one ξ to another. This implies that the Hall coefficient, H_k is quantised for low energies:

$$H_k = \frac{B_k}{ne} = \frac{h}{\xi e^2} \quad (5.211)$$

this is because $n_e = \frac{\xi N_p}{L_x L_y}$ for Landau levels, thus the transition from one step to another corresponds to the filling of a new Landau level.

This is in fact the case: every time the energy of a Landau level approaches the fermi energy it is possible for electrons to scatter across the sample, from the edge states that are going in the one direction to the edge states that are going in the other direction. The Landau levels are inert, so conduction only occurs in the edge states where the energy bands curl upwards and *cross* the fermi level.

The scattering *across* the sample causes a current drop, giving a higher on-diagonal resistivity.

GENERAL RELATIVITY & COSMOLOGY

6.1 SPECIAL RELATIVITY

Spacetime consists of events whose coordinates x^μ are four-vectors:

$$x^\mu = (ct, x^1, x^2, x^3) \quad (6.1)$$

however we set $c = 1$ which means that time is measured in units of length. The Minkowski metric is the metric of free space, which can in a fitting coordinate system be written as

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6.2)$$

Using this metric we define the proper length:

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \quad (6.3)$$

which is invariant under

- *Translations*

$$x^\mu \rightsquigarrow x^\mu + a^\mu \quad (6.4)$$

- *Rotations*

$$x^\mu \rightsquigarrow \Lambda^\mu_\nu x^\nu, \quad \Lambda^\mu_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6.5)$$

i.e rotations about 0 – *i*-axes

- *Boosts*

$$x^\mu \rightsquigarrow \Lambda^\mu_\nu x^\nu, \quad \Lambda^\mu_\nu = \begin{pmatrix} \cosh \phi & -\sinh \phi & 0 & 0 \\ -\sinh \phi & \cosh \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6.6)$$

where the boost parametre, ϕ , is related to the velocity difference between the initial and final inertial systems:

$$\tanh \phi = v \quad (6.7)$$

which allows us to write this in a more familiar way:

$$\Lambda_{\nu}^{\mu} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \gamma \equiv (1 - v^2)^{-\frac{1}{2}} \quad (6.8)$$

We differentiate between line-segments depending on the sign of ds^2 :

- *Spacelike separation*: $ds > 0$
- *Null separation*: $ds = 0$
- *Timelike separation*: $ds < 0$

For time like curves we can define the proper time:

$$d\tau^2 = -ds^2 = -\eta_{\mu\nu} dx^{\mu} dx^{\nu} \quad (6.9)$$

Which allows us to define the relativistic velocity:

$$u^{\mu} = \frac{dx^{\mu}}{d\tau} = \gamma(1, v^1, v^2, v^3) \quad (6.10)$$

Note that

$$u_{\mu} u^{\mu} = \frac{1}{1 - v^2} (v^2 - 1) = -1 \quad (6.11)$$

we can also define the relativistic acceleration:

$$a^{\mu} = \frac{du^{\mu}}{d\tau} = \frac{d^2 x^{\mu}}{d\tau^2} \quad (6.12)$$

Curves

We have up till now dealt with infinitesimal line-segments, we will now generalise proper lengths and proper times to curves. Given a curve $x^{\mu}(\lambda)$:

$$\Delta\tau = \int_{\lambda_1}^{\lambda_2} d\lambda \sqrt{-\eta_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda}} \quad (6.13)$$

6.2 EQUIVALENCE PRINCIPLE

EINSTEIN'S EQUIVALENCE PRINCIPLE:

In small enough regions of spacetime the laws of physics reduce to those of special relativity. Therefore, it is impossible to measure the presence of gravity through local experiments.

Due to this equivalence principle we choose to measure acceleration with respect to *freely falling particles*.

6.2.1 Gravitational Redshift

Before we introduce the language of general relativity, we can already show a physical effect due to the principle of equivalence, which contradicts the predictions made by Newtonian physics:

Consider a tower of height h , where a laser at the bottom of the tower emits light upwards and the wavelength of the light is measured at the top. Due to the fact that both the top and the bottom of the tower are accelerating upwards the wavelength at the bottom of the tower will be different from that at the top, due to the finite speed of light:

$$\Delta\lambda = \lambda_{\text{top}} - \lambda_{\text{bottom}} = \frac{a_g h}{c^2} \lambda_{\text{bottom}} \quad (6.14)$$

Generally we can express this as

$$\frac{\Delta\lambda}{\lambda} = \frac{1}{c^2} \Delta\phi \quad (6.15)$$

where $\Delta\phi$ is the potential difference between the two points.

This also has the implication the gravitational fields dilate time:

$$cT_{\text{bottom}} = \lambda_{\text{bottom}}, \quad cT_{\text{top}} = \lambda_{\text{top}} \quad (6.16)$$

hence

$$\Delta T = \frac{\Delta\phi}{c^2} T_{\text{bottom}}, \quad \frac{T_{\text{top}}}{T_{\text{bottom}}} = 1 + \frac{\Delta\phi}{c^2} \quad (6.17)$$

which tells us that clocks go slower inside gravitational wells.

6.3 GENERAL SPACETIMES

The Minkowski metric is only valid when there is no gravity (except for local inertial systems, that is). As we have seen gravity influences the passage of time, and it also influence relative lengths, therefore we can see that we need a more general metric than the Minkowski metric: $\eta_{\mu\nu} \rightsquigarrow g_{\mu\nu}$. Hence

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (6.18)$$

$g_{\mu\nu}$ is a tensor, therefore it transforms as a tensor:

TRANSFORMATION OF THE METRIC

Suppose we are given a (bijective) coordinate transform:

$$\tilde{x}^\mu = \tilde{x}^\mu(x^\nu) \quad (6.19)$$

Then the metric, which is a $(0, 2)$ -tensor, transforms like

$$\tilde{g}_{\alpha\beta} = \frac{dx^\mu}{d\tilde{x}^\alpha} \frac{dx^\nu}{d\tilde{x}^\beta} g_{\mu\nu} \quad (6.20)$$

The inverse metric $g^{\mu\nu}$ is defined as the tensor that satisfies

$$g_{\mu\nu}g^{\nu\rho} = \delta_\mu^\rho \quad (6.21)$$

which transforms like a $(2,0)$ -tensor:

TRANSFORMATION OF THE INVERSE-METRIC

Suppose we are given a (bijective) coordinate transform:

$$\tilde{x}^\mu = \tilde{x}^\mu(x^\nu) \quad (6.22)$$

Then the inverse metric, which is a $(2,0)$ -tensor, transforms like

$$\tilde{g}^{\alpha\beta} = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \frac{\partial \tilde{x}^\beta}{\partial x^\nu} g^{\mu\nu} \quad (6.23)$$

Geodesics

Freely falling particles follow curves that maximise proper time; these paths are called geodesics. Using this principle we can, using the calculus of variations find a differential equation that describes geodesics.

We want to find the curve, $x^\mu(\lambda)$ that minimises:

$$\Delta\tau[x^\mu(\lambda)] = \int_{\lambda_1}^{\lambda_2} d\lambda \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \quad (6.24)$$

hence we require $\delta\Delta\tau = 0$, which results in the following equation

GEODESIC EQUATION

The motion of a freely particle is described by a curve, $x^\mu(\tau)$, which satisfies:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = 0 \quad (6.25)$$

where the Christoffel Symbol is defined by

$$\Gamma_{\alpha\beta}^\mu = \frac{1}{2}g^{\mu\sigma} (\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}) \quad (6.26)$$

*The Christoffel Symbol is
not a tensor!*

The Christoffel Symbol is symmetric in its lower indices.

Newtonian Limit

When we look at very weak, time-independent gravity ($g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}(x)$ for $h_{\mu\nu}$ small), and we require low velocities ($|\frac{dx^i}{d\tau}| \ll 1$) general relativity should simplify to Newtons laws of gravitation. In this limit the geodesic equation is

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma_{00}^\mu \left(\frac{dt}{d\tau} \right)^2 = 0 \quad (6.27)$$

Up to linear order in $h_{\mu\nu}$ this becomes

$$\frac{d^2\mathbf{x}}{d\tau^2} = \frac{1}{2}\nabla h_{00} \quad (6.28)$$

hence we can interpret the time-time entry in the metric as the gravitational potential in the Newtonian limit:

$$g_{00} \approx -1 - 2\phi, \quad \text{Newtonian Limit} \quad (6.29)$$

6.3.1 Local Inertial System

LOCAL INERTIAL SYSTEM

Consider an event, p . There always exists a coordinate system in which the metric simplifies to the Minkowski metric as per Equation 6.20. This coordinate system is referred to as a *Local Inertial System*, for which we have

$$g_{\mu\nu}|_p = \eta_{\mu\nu}, \quad \partial_\rho g_{\mu\nu}|_p = 0 \quad (6.30)$$

This implies that in the local inertial system $\Gamma_{\mu\nu}^\rho = 0$, which ties back to the implication of Einstein's Equivalence Principle, that we cannot detect the presence of gravity *locally*.

6.4 TENSORS AND THE PRINCIPLE OF COVARIANCE

The simplest tensor is a tensor with zero indices, a scalar. Scalars transform rather trivially, consider $\tilde{x}^\mu = \tilde{x}^\mu(x^\mu)$:

$$\tilde{\Phi}(\tilde{x}^\mu) = \Phi(x^\mu) \quad (6.31)$$

The next simplest tensor has one index, which we can either place high or low. If it is placed high the vector is a contravariant vector, V^μ , which we also just call a vector. If we place the index down low it is a covariant vector, or a dual vector, V_μ . These vectors are related to each other:

$$V_\mu = g_{\mu\nu} V^\nu \quad (6.32)$$

These quantities transform quite differently from scalar fields:

$$\tilde{V}_\mu = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} V_\nu, \quad \tilde{V}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} V^\nu \quad (6.33)$$

Generally though

TENSOR TRANSFORMATION

Given a (m, n) -tensor $T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n}$ it transforms as

$$\tilde{T}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n} = \prod_{(i,j)=(1,1)}^{(m,n)} \frac{\partial \tilde{x}^{\alpha_i}}{\partial x^{\mu_i}} \frac{\partial x^{\nu_j}}{\partial \tilde{x}^{\beta_j}} T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n} \quad (6.34)$$

6.4.1 Covariant Derivative

We need to generalise the concept of a derivative, so that we can express our equations in a covariant form; in a form that is independent of coordinate system. This is very easy for scalar fields, because their derivatives are already covariant:

$$\tilde{\partial}_\mu \tilde{\Phi} = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \partial_\nu \Phi \quad (6.35)$$

however, this is not the case generally. For instance, let us look at how the derivative of a vector field, $\partial_\mu V^\nu$, transforms:

$$\tilde{\partial}_\alpha \tilde{V}^\beta = \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} \partial_\mu \left(\frac{\partial \tilde{x}^\beta}{\partial x^\nu} V^\nu \right) \quad (6.36)$$

$$= \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} \frac{\partial \tilde{x}^\beta}{\partial x^\nu} \partial_\mu V^\nu + \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} \frac{\partial^2 \tilde{x}^\beta}{\partial x^\mu \partial x^\nu} V^\nu \quad (6.37)$$

$$= \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} \frac{\partial \tilde{x}^\beta}{\partial x^\nu} \left(\partial_\mu V^\nu + \frac{\partial x^\nu}{\partial \tilde{x}^\gamma} \frac{\partial^2 \tilde{x}^\gamma}{\partial x^\mu \partial x^\rho} V^\rho \right) \quad (6.38)$$

The second term is the problem, if this term were equal to zero we would have shown that the partial derivative of a vector field is covariant. Let us define the covariant derivative, D_μ , which is the generalisation of the partial derivative, which reduces to the partial derivative in a local inertial system. In a local inertial system it can be shown that

$$\Gamma_{\mu\rho}^\nu|_{\text{LIS},p} = \frac{\partial x^\nu}{\partial \tilde{x}^\gamma} \frac{\partial^2 \tilde{x}^\gamma}{\partial x^\mu \partial x^\rho} \quad (6.39)$$

We are now ready to define the covariant derivative of vector field:

COVARIANT DERIVATIVE OF A VECTOR-FIELD

Given a general spacetime, $g_{\mu\nu}$, and a vector field, V^ρ , the covariant derivative is defined by

$$D_\mu V^\rho = \partial_\mu V^\rho + \Gamma_{\mu\sigma}^\rho V^\sigma \quad (6.40)$$

where $\Gamma_{\mu\sigma}^\rho$ is the Christoffel Symbol defined in Equation 6.26

The covariant derivative of a dual-vector is almost the same

COVARIANT DERIVATIVE OF A DUAL-VECTOR-FIELD

Given a general spacetime, $g_{\mu\nu}$, and a vector field, V_ρ , the covariant derivative is defined by

$$D_\mu V_\rho = \partial_\mu V_\rho - \Gamma_{\mu\rho}^\sigma V_\sigma \quad (6.41)$$

where $\Gamma_{\mu\sigma}^\rho$ is the Christoffel Symbol defined in Equation 6.26

The covariant derivative of a general (m, n) -tensor is

$$\begin{aligned} D_\rho T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n} &= \partial_\rho T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n} + \Gamma_{\rho\sigma}^{\mu_1} T^{\sigma \mu_2 \dots \mu_m}_{\nu_1 \dots \nu_n} \\ &\quad + \Gamma_{\rho\sigma}^{\mu_2} T^{\mu_1 \sigma \mu_3 \dots \mu_m}_{\nu_1 \dots \nu_n} + \dots + \Gamma_{\rho\sigma}^{\mu_m} T^{\mu_1 \dots \mu_{m-1} \sigma}_{\nu_1 \dots \nu_n} \\ &\quad - \Gamma_{\rho\nu_1}^\sigma T^{\mu_1 \dots \mu_m}_{\sigma \nu_2 \dots \nu_n} - \Gamma_{\rho\nu_2}^\sigma T^{\mu_1 \dots \mu_m}_{\nu_1 \sigma \nu_3 \dots \nu_n} \\ &\quad - \dots - \Gamma_{\rho\nu_n}^\sigma T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_{n-1} \sigma} \end{aligned} \quad (6.42)$$

Covariant derivative along a curve

Consider a curve $x^\mu(\lambda)$ then we can calculate the covariant derivative along this curve. For instance the covariant of a vector along the curve would be

$$\frac{D}{d\lambda} V^\mu \equiv \frac{dx^\rho}{d\lambda} D_\rho V^\mu \quad (6.43)$$

If the covariant derivative along a curve of a quantity is zero, then the quantity is conserved along the curve. For example if V^μ is parallel to u^μ at the beginning of the curve and the quantity $V_\mu u^\mu$ is conserved along the curve, then we can conclude that the vectors remain parallel along the curve.

Properties of the covariant derivative

The covariant derivative, like other derivatives we have

- linearity
- a product rule

COVARIANT DERIVATIVE OF THE METRIC

Given a general spacetime with metric $g_{\mu\nu}$ and inverse metric $g^{\mu\nu}$, then we have

$$D_\rho g_{\mu\nu} = D_\rho g^{\mu\nu} = 0 \quad (6.44)$$

which can easily be shown by transforming into a local inertial system

6.4.2 Acceleration in General Relativity

In the special relativity section we defined the four-velocity, $u^\mu = \frac{dx^\mu}{d\tau}$. We now define the covariant acceleration:

$$a^\mu = \frac{D}{d\tau} u^\mu = \frac{dx^\nu}{d\tau} D_\nu u^\mu \quad (6.45)$$

Writing this out, using the definition of the covariant derivative:

$$a^\mu = \frac{dx^\nu}{d\tau} (\partial_\nu u^\mu + \Gamma_{\nu\rho}^\mu u^\rho) \quad (6.46)$$

$$= \frac{dx^\nu}{d\tau} \frac{\partial u^\mu}{\partial x^\nu} + \Gamma_{\nu\rho}^\mu \frac{dx^\nu}{d\tau} u^\rho \quad (6.47)$$

$$= \frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\nu\rho}^\mu \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau} \quad (6.48)$$

This simplifies to the geodesic equation in the case where $a^\mu = 0$, i.e. when there are no non-gravitational forces. Additionally, in a local inertial system $a^\mu = \frac{d^2 x^\mu}{d\tau^2}$, which is the acceleration we defined for special relativity, as we would expect.

For massive particles travelling on geodesics we have

$$a^\mu = 0, \quad u_\mu u^\mu = -1 \quad (6.49)$$

however, for null-curves (for which $d\tau = 0$) we instead have

$$a^\mu = 0, \quad g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0 \quad (6.50)$$

where λ is the *affine parametre* we use to parametrise the null curve.

NB: τ cannot be used to parametrise the null curve, because it is constant along the curve.

6.4.3 Maxwell's Equations in a general spacetime

Maxwell's equations in vacuum can be expressed as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad D_\mu F^{\mu\nu} = -J^\nu \quad (6.51)$$

We can see that the electromagnetic field strength tensor, $F_{\mu\nu}$ is antisymmetric. These equations are covariant, which is easy to see for the second equation, however the appearance of partial derivatives in the first equation makes it look like this may not be covariant. However, due to the symmetry of the Christoffel Symbols in its lower indices we have

$$D_\mu A_\nu - D_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (6.52)$$

Assuming a particle is only subject to the electromagnetic force then its covariant acceleration can be calculated through

$$a^\mu = -\frac{q}{m} g_{\nu\rho} u^\nu F^{\rho\mu} \quad (6.53)$$

6.5 RIEMANN CURVATURE TENSOR

We are looking for an object that describes the curvature of the metric, $g_{\mu\nu}$, and hence of spacetime in a manner that all coordinate systems agree upon: we are looking for a covariant expression of the curvature. Unfortunately our best candidate, $\Gamma_{\mu\nu}^\rho$, cannot be used, because it is not a tensor

It does not transform like a tensor. This is easy to show, $\Gamma_{\mu\nu}^\rho = 0$ in a local inertial system, but nonzero generally which cannot hold for a tensor, see Equation 6.34.

We instead will look at a derivative of the Christoffel Symbol, but we wish for a covariant expression.

We cannot just use $D_\sigma \Gamma_{\mu\nu}^\rho$ because this is not a tensor either

Consider an infinitesimal square with (infinitesimal) width w^μ and (infinitesimal) height h^μ , with starting position x_0^μ . We wish to transport a vector V^μ along the curve in such a way that the covariant derivative is and remains zero, this process is referred to as parallel transport. We, therefore, demand that

$$dV^\mu = -\Gamma_{\rho\nu}^\mu V^\nu dx^\rho \quad (6.54)$$

We will ignore terms of higher order than h^2 (and hence w^2 and hw)

Let us write this in a more legible way

$$V^\mu(x_0 + w) - V^\mu(x_0) = -\Gamma_{\rho\nu}^\mu(x_0) V^\nu(x_0) w^\rho \quad (6.55)$$

going to the next corner

$$\begin{aligned}\Delta V_1 &= V^\mu(x_0 + w + h) - V^\mu(x_0) = -\Gamma_{\rho\nu}^\mu(x_0)V^\nu(x_0)w^\rho \\ &\quad - \Gamma_{\rho\nu}^\mu(x_0 + w)V^\nu(x_0 + w)h^\rho\end{aligned}\quad (6.56)$$

But we can write this out

$$\begin{aligned}\Delta V_1 &= -\Gamma_{\rho\nu}^\mu(x_0)V^\nu(x_0)w^\rho \\ &\quad - (\Gamma_{\rho\nu}^\mu(x_0) + \partial_\alpha \Gamma_{\rho\nu}^\mu(x_0)w^\alpha) (V^\nu(x_0) - \Gamma_{\alpha\beta}^\nu V^\alpha(x_0)w^\beta) h^\rho\end{aligned}\quad (6.57)$$

we can also get there the other way, so first adding h and then w :

$$\begin{aligned}\Delta V_2 &= -\Gamma_{\rho\nu}^\mu V^\nu h^\rho \\ &\quad - (\Gamma_{\rho\nu}^\mu + \partial_\alpha \Gamma_{\rho\nu}^\mu h^\alpha) (V^\nu - \Gamma_{\alpha\beta}^\nu V^\alpha h^\beta) w^\rho\end{aligned}\quad (6.58)$$

There difference $\Delta V = \Delta V_1 - \Delta V_2$ is also the difference between to the two parallel transported vectors (along the different paths). By renaming some of the dummy variables we end up with

$$\Delta V^\mu = (\partial_\rho \Gamma_{\sigma\nu}^\mu - \partial_\sigma \Gamma_{\rho\nu}^\mu - \Gamma_{\sigma\alpha}^\mu \Gamma_{\nu\rho}^\alpha + \Gamma_{\rho\alpha}^\mu \Gamma_{\nu\sigma}^\alpha) \quad (6.59)$$

RIEMANN CURVATURE TENSOR

The Riemann Curvature Tensor is defined as

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma_{\nu\sigma}^\rho - \partial_\nu \Gamma_{\mu\sigma}^\rho + \Gamma_{\mu\alpha}^\rho \Gamma_{\sigma\nu}^\alpha - \Gamma_{\nu\alpha}^\rho \Gamma_{\sigma\mu}^\alpha \quad (6.60)$$

Clearly $R^\rho_{\sigma\mu\nu} = 0$ if and only if we are in Minkowski Space.

6.5.1 Symmetries of the Riemann Curvature Tensor

$R^\rho_{\sigma\mu\nu}$ is antisymmetric in μ and ν as swapping these indices would correspond to switching the two paths discussed in the derivation. To discuss further symmetries let us lower the upper index:

$$R_{\mu\nu\rho\sigma} = g_{\mu\alpha} R^\alpha_{\nu\rho\sigma} \quad (6.61)$$

we have that

$$R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma} = -R_{\mu\nu\sigma\rho} \quad (6.62)$$

and

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu} \quad (6.63)$$

and finally

$$R_{\mu\nu\rho\sigma} + R_{\mu\sigma\nu\rho} + R_{\mu\rho\sigma\nu} = 0 \quad (6.64)$$

Ricci Tensor and Scalar

We can contract the first and third indices of the Riemann tensor, to obtain the Ricci tensor:

$$R_{\mu\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu} = R^\alpha_{\mu\alpha\nu} \quad (6.65)$$

which is a symmetric tensor

The trace of the Ricci tensor is the Ricci scalar:

$$g^{\mu\nu} R_{\mu\nu} = R = R^\alpha_{\alpha} \quad (6.66)$$

6.5.2 Bianchi Identity

The following identity holds for the Riemann curvature tensor:

$$D_\alpha R_{\mu\nu\rho\sigma} + D_\nu R_{\alpha\mu\rho\sigma} + D_\mu R_{\nu\alpha\rho\sigma} = 0 \quad (6.67)$$

Due to the fact that the covariant derivative of the metric is zero we can freely contract indices inside the Bianchi Identity:

$$0 = g^{\nu\sigma} g^{\mu\rho} (D_\alpha R_{\mu\nu\rho\sigma} + D_\nu R_{\alpha\mu\rho\sigma} + D_\mu R_{\nu\alpha\rho\sigma}) \quad (6.68)$$

$$= D_\alpha R - 2D^\sigma R_{\sigma\alpha} \quad (6.69)$$

Therefore, for the Ricci tensor we can compress (some of?) the information from the Bianchi Identity into a simpler equation:

$$D^\mu \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) = 0 \quad (6.70)$$

6.6 EINSTEIN'S EQUATIONS

We've already seen that the metric is the general relativistic counterpart of the Newtonian (gravitational) potential, therefore we might be inclined to assume that the Riemann Curvature Tensor is the general relativistic counterpart of the laplacian of the potential. This is indeed the case, however, as we will see it suffices to use the Ricci Tensor.

However, we know need to find the general relativistic counterpart of the mass-density. This object is clearly a $(0, 2)$ -tensor, like the metric, and in order for the Newtonian limit to make any sense we can already state that T_{00} will simplify to the mass-density in the Newtonian limit. This object is defined below

ENERGY-MOMENTUM TENSOR

The energy momentum tensor, $T^{\mu\nu}(x) \prod_{\rho=0, \rho \neq \nu} dx^\rho$, is the flux of the μ^{th} component of the four-momentum across an infinitesimal 3-surface $\prod_{\rho=0, \rho \neq \nu} dx^\rho$.

This definition is difficult to interpret, so let us write this more concretely

$$\begin{aligned} T^{00} &: \text{Energy density} \\ T^{i0} &: \text{Density of the } i^{\text{th}} \text{ component of the momentum} \\ T^{0j} &: \text{Energy flux through the surface orthogonal to } x^j \\ T^{ij} &: \text{Internal forces per unit area (pressure and shear)} \end{aligned}$$

Energy flux is the same as momentum density ($c = 1$), and similarly for T^{ij} and T^{ji} . Therefore, the energy-momentum tensor is symmetric

$$T^{\mu\nu} = T^{\nu\mu} \quad (6.71)$$

This is because everyday pressures (like 1atm) are $\sim 10^{-12} c^2 \rho_m$, hence we can readily neglect these

In the Newtonian limit we already know what happens:

$$T^{\mu\nu} = \begin{pmatrix} \rho_m & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (6.72)$$

we call this kind of matter *Newtonian Matter*. A perfect fluid (zero viscosity and zero heat conduction) has the following energy-momentum tensor:

$$T_{\mu\nu} = (\rho + p)u^\mu u^\nu + p\eta^{\mu\nu} \quad (6.73)$$

The specific type of fluid can be expressed through the function $p = p(\rho)$:

$$\begin{aligned} \text{Dust:} \quad & p = 0 \\ \text{Photons:} \quad & p = \frac{1}{3}\rho \end{aligned}$$

6.6.1 Conservation of the Energy-Momentum Tensor

In special relativity we have the identity

$$\partial_\mu T^{\mu\nu} = 0 \quad (6.74)$$

this is a continuity equation: the (temporal) change in energy is due to energy flux out of a volume element.

We can quite easily generalise this to a covariant equation:

$$D_\mu T^{\mu\nu} = 0 \quad (6.75)$$

indeed, the energy momentum tensor is conserved. The metric is also a conserved quantity, so we can raise and lower indices which implies

$$D^\mu T_{\mu\nu} = 0 \quad (6.76)$$

6.6.2 Einstein's Equations

We know have two conserved quantities, $T_{\mu\nu}$ and $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$, thus these quantities must be equal up to an "integration" constant:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \quad (6.77)$$

strictly speaking it is not an integration constant, just a tensor that also whose covariant derivative is zero

where Λ is a free parametre, which later leads to the concept of dark energy and the expansion of the universe. The Λ -term can be thought of as a zero-point energy, i.e even when $T_{\mu\nu} = 0$ there is energy in the universe, which is why I think it would make sense to write the term on the right hand side. We can show that the $8\pi G$ term must be there if we want to use our prior definition of $T_{\mu\nu}$; it appears when we take the Newtonian limit.

6.7 SCHWARZSCHILD METRIC

Let us consider a spherically symmetric mass distribution. This symmetry must also apply to the metric, therefore we are interested in a line-segment that is spherically symmetric. Beginning with the most general line-segment in spherical coordinates:

$$ds^2 = -A(t, r)dt^2 + B(t, r)dr^2 + 2C(t, r)dt dr + D(t, r)r^2 d\Omega^2 \quad (6.78)$$

the functions A, B, C and D can only depend on time and r due to the aforementioned spherical symmetry. We can diagonalise the (t, r) -part by changing basis, and by redefining r we can also remove D . We additionally write g_{tt} and g_{rr} using exponentials, to simplify our calculations later:

$$ds^2 = -e^{2\alpha(r)} dt^2 + e^{2\beta(r)} dr^2 + r^2 d\Omega^2 \quad (6.79)$$

We have also removed the time dependence; even with time-dependence we will obtain the same result, as stated by *Birkhoff's Theorem*. By plugging this metric into Einstein's Equation and assuming that $T_{\mu\nu} = 0$ outside the spherically symmetric object we get differential equations that describe the functions α and β :

$$\begin{aligned} R_{tt} &= e^{2(\alpha-\beta)} \left(\partial_r^2 \alpha + (\partial_r \alpha)^2 - \partial_r \alpha \partial_r \beta + \frac{2}{r} \partial_r \alpha \right) = 0 \\ R_{rr} &= -\partial_r^2 \alpha + \partial_r \alpha \partial_r \beta - (\partial_r \alpha)^2 + \frac{2}{r} \partial_r \beta = 0 \\ R_{\theta\theta} &= \frac{1}{\sin^2 \theta} R_{\phi\phi} = 1 + e^{-2\beta} (r(\partial_r \beta - \partial_r \alpha) - 1) \end{aligned}$$

We can take linear combinations of the elements of the Ricci tensor, giving us

$$\partial_r (\alpha + \beta) = 0 \quad (6.80)$$

we can absorb the integration constant into our definition of t , so we can conclude that $\alpha = -\beta$. The final differential equation we need to solve is thus

$$1 - e^{2\alpha} (1 + 2r \partial_r \alpha) = 0 \quad (6.81)$$

let $2\alpha = \ln(f)$:

$$f + r \partial_r f = 1 = \partial_r (rf) \quad (6.82)$$

hence

$$f = 1 - \frac{r_0}{r} \quad (6.83)$$

in conclusion

$$g_{\mu\nu} = \begin{pmatrix} -(1 - \frac{r_0}{r}) & 0 & 0 & 0 \\ 0 & (1 - \frac{r_0}{r})^{-1} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad (6.84)$$

By once again looking at the Newtonian limit we can conclude $r_0 = 2GM$, the *Schwarzschild Radius*.

6.7.1 Geodesics

Define

$$\dot{x}^\mu \equiv \frac{dx^\mu}{d\tau}, \quad \ddot{x}^\mu \equiv \frac{d\dot{x}^\mu}{d\tau} \quad (6.85)$$

This makes the geodesic equation a bit more palatable:

$$\ddot{x}^\mu + \Gamma_{\alpha\beta}^\mu \dot{x}^\alpha \dot{x}^\beta = 0, \quad g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = -1 \quad (6.86)$$

for massive particles following time-like curves. Consider now the Christoffel Symbols:

$$\Gamma_{tr}^t = \frac{1}{2} \partial_r \ln \left(1 - \frac{r_0}{r} \right), \quad \Gamma_{r\theta}^\theta = \Gamma_{r\phi}^\phi = \frac{1}{r} \quad (6.87)$$

$$\Gamma_{\phi\phi}^\theta = -\sin \theta \cos \theta, \quad \Gamma_{\theta\phi}^\phi = \cot \theta \quad (6.88)$$

There are other nonzero components, however, we will only use the aforementioned in the following. Using these we find that

$$\ddot{t} + \partial_r \ln \left(1 - \frac{r_0}{r} \right) \dot{t} \dot{r} = 0 \quad (6.89)$$

$$\ddot{\theta} + \frac{2}{r} \dot{r} \dot{\theta} - \sin \theta \cos \theta \dot{\phi}^2 = 0 \quad (6.90)$$

$$\ddot{\phi} + \frac{2}{r} \dot{r} \dot{\phi} + 2 \cot \theta \dot{\theta} \dot{\phi} = 0 \quad (6.91)$$

we can write these slightly differently though

$$\frac{d}{d\tau} \left[\dot{t} \left(1 - \frac{r_0}{r} \right) \right] = 0 \quad (6.92)$$

$$\frac{d}{d\tau} (r^2 \dot{\theta}) = r^2 \sin \theta \cos \theta \dot{\phi}^2 \quad (6.93)$$

$$\frac{d}{d\tau} (r^2 \sin^2 \theta \dot{\phi}) = 0 \quad (6.94)$$

and for the r component we instead use the formula for a line segment and use the normalisation of \dot{x}^μ :

$$-1 = - \left(1 - \frac{r_0}{r} \right) \dot{t}^2 + \frac{\dot{r}^2}{1 - \frac{r_0}{r}} + r^2 (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \quad (6.95)$$

Constricting ourselves to $\theta = \frac{\pi}{2}$ which we can always do by rotating coordinate systems. Under this choice of coordinates we get

the angular momentum vector is still a conserved quantity, as we will see

$$\frac{d}{d\tau} \left[\dot{t} \left(1 - \frac{r_0}{r} \right) \right] = 0 \quad (6.96)$$

$$\frac{d}{d\tau} (r^2 \dot{\phi}) = 0 \quad (6.97)$$

$$-1 = - \left(1 - \frac{r_0}{r} \right) \dot{t}^2 + \frac{\dot{r}^2}{1 - \frac{r_0}{r}} + r^2 \dot{\phi}^2 \quad (6.98)$$

The first two equations describe conserved quantities, the first being the energy and the second the angular momentum.

$$E = \dot{t} \left(1 - \frac{r_0}{r} \right) \quad (6.99)$$

$$L = r^2 \dot{\phi} \quad (6.100)$$

using which we can simplify the third equation significantly

$$E^2 = 1 - \frac{2GM}{r} + \dot{r}^2 + \frac{L^2}{r^2} \left(1 - \frac{2GM}{r} \right) \quad (6.101)$$

This differs from the classical counterpart only by the final term, that goes like r^3 . We can change coordinate systems to $X = \frac{GM}{r}$ and express this as a differential equation in ϕ :

$$\frac{d^2 X}{d\phi^2} + X - 3X^2 = \left(\frac{GM}{L} \right)^2 \quad (6.102)$$

here it is the $3X^2$ that is the relativistic correction to Newton's equation. Define $Y = \left(\frac{L}{GM} \right)^2 X$ we get

$$\frac{d^2 Y}{d\phi^2} + Y - 3 \left(\frac{GM}{L} \right)^2 Y^2 = 1 \quad (6.103)$$

hence the Newtonian limit clearly corresponds to $X \ll \left(\frac{GM}{L} \right)^2$, which is either achieved by looking at large values of r or small values of L .

$$\frac{1}{r} \ll \frac{GM}{L^2} \quad (6.104)$$

In the intermediate limit, so where we can treat the Newtonian solution as the solution the unperturbed problem, and let the Y^2 term perturb the solution, we begin with

$$r(\phi) = \frac{(1 - e^2)a}{1 + e \cos \phi} \quad (6.105)$$

where

$$E^2 = 1 + \left(\frac{GM}{L} \right)^2 (e^2 - 1) \quad \frac{L^2}{GM} = (1 - e^2)a \quad (6.106)$$

e is the eccentricity of the orbit, and a is its semimajor axis. Thus we look at

$$X_0(\phi) = \left(\frac{GM}{L} \right)^2 (1 + e \cos \phi) \quad (6.107)$$

and calculate the first order correction

$$\frac{d^2 X_1}{d\phi^2} + X_1 = 3X_0^2 \quad (6.108)$$

giving us

$$X_1(\phi) = \left(\frac{GM}{L} \right)^4 \left(3 \left(1 + \frac{e^2}{2} + e\phi \sin \phi \right) - \frac{e^2}{2} \cos 2\phi \right) \quad (6.109)$$

Actually, the largest part of Mercury's precession is accounted for classically, this just accounts for a small deviation from the classical precession.

the solution is no longer periodic, which is what explain Mercury's precession. Using this we can calculate the precession of an elliptic per period:

$$\Delta\phi = \frac{6\pi GM}{(1 - e^2)ac^2} \quad (6.110)$$

6.7.2 Deflection of Light

The geodesic equation for null-curves also predicts things that aren't found classically, as we will see:

$$\ddot{x}^\mu + \Gamma_{\alpha\beta}^\mu \dot{x}^\alpha \dot{x}^\beta = 0, \quad g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = 0 \quad (6.111)$$

however now we cannot use τ as a parametre any more, so for null curves we need

$$\dot{x}^\mu = \frac{dx^\mu}{d\lambda}, \quad \ddot{x}^\mu = \frac{d\dot{x}^\mu}{d\lambda} \quad (6.112)$$

The fact that these are null-curves actually doesn't change that much. We still have energy and angular momentum conservation:

$$E = \frac{dt}{d\lambda} \left(1 - \frac{r_0}{r}\right), \quad L = r^2 \frac{d\phi}{d\lambda} \quad (6.113)$$

and the differential equation describing X becomes

$$\frac{d^2 X}{d\phi^2} + X = 3X^2 \quad (6.114)$$

so the only difference is that we no longer have the $\left(\frac{GM}{L}\right)^2$ term. Notice that the differential equation in the absence of X^2 describes straight lines, as we might expect. However the X^2 term curves the straight lines. The deflection angle for a lightray with impact parametre, b , due to GR is

$$\Delta\phi = \frac{4GM}{bc^2} \quad (6.115)$$

6.8 BLACK HOLES

6.8.1 Schwarzschild Metric

We know that energy is conserved, which tells us that

$$d\tau = \frac{1 - \frac{r_0}{r}}{E} dt \quad (6.116)$$

Thus at $r = r_0$ the proper time stops, with respect to t . Let us require that for $r \rightsquigarrow \infty$ the proper time and coordinate times agree with each other, which implies that the infalling object is at rest at infinity: $E = 1$. Thus

$$\frac{dr}{d\tau} = -\sqrt{\frac{r_0}{r}} \rightsquigarrow r(\tau)^{\frac{3}{2}} - r(\tau_0)^{\frac{3}{2}} = \frac{3}{2} \sqrt{r_0} (\tau_0 - \tau) \quad (6.117)$$

The distant observer measures t , and when comparing to τ the distant observer claims that time stops at the Schwarzschild metric

Thus according to the in-falling observer nothing special happens at r_0 ; the only interesting point is at $r = 0$, which is where the velocity diverges. However, according to the distant observer we get

$$\frac{dr}{dt} = \frac{dt}{d\tau} \frac{dr}{d\tau} = -\left(1 - \frac{r_0}{r}\right) \sqrt{\frac{r_0}{r}} \quad (6.118)$$

close to $r = r_0$ we can approximate this as

$$\frac{dr}{dt} \approx -\frac{r - r_0}{r_0} \quad (6.119)$$

which is an exponential decay

$$r(t) \approx r_0 + r_C \exp\left(-\frac{t}{r_0}\right) \quad (6.120)$$

hence according to the distant observer the event horizon is never reached.

Closing of Light Cones

In spacetime diagrams in special relativity we always know that massive objects' motions are limited by the light cones. The same holds in general relativity, however, here the light cones are more complicated. In the following we look at radial null curves, which will limit the motions of massive objects that also are moving radially. Radial null-curves have

$$\frac{dr}{dt} = \pm \left(1 - \frac{r_0}{r}\right) \quad (6.121)$$

according to the distant observer. At $r \approx r_0$ we see that the apparent speed of light is zero, which again corresponds to the proper time stopping according to the distant observer. This implies that the infalling and outgoing null-curves both have zero (radial) velocity at the Schwarzschild radius, therefore no massive object can enter or leave the black hole, according to the distant observer.

Eddington-Finkelstein Coordinates

As we have seen, the fact that time stops at the event horizon is coordinate dependent. This is because the metric has a singularity at $r = r_0$ (and one at $r = 0$). The singularity at $r = r_0$ is an artefact of the coordinate system, *not* of the physics, therefore, there exists coordinate systems where this singularity is not present. However, the singularity at $r = 0$ cannot be mended with coordinate transforms; it is a singularity in the spacetime.

Let us now define a coordinate system that mends the $r = r_0$ singularity, by redefining the radial component

$$r_* = r + r_0 \ln \left| \frac{r}{r_0} - 1 \right| \quad (6.122)$$

clearly

$$dr_* = \frac{dr}{1 - \frac{r_0}{r}} \quad (6.123)$$

In these coordinates radial null curves obey

$$\frac{dr_*}{dt} = \pm 1 \quad \rightsquigarrow \quad r_*(t) = \pm t + r_*(0) \quad (6.124)$$

thus using this as a radial coordinate will fix the problem we just saw with light cones closing. However the function $r(r_*)$ is not particularly pretty, therefore, it is beneficial to work a bit harder for our new coordinate system. Define

$$u = t - r^*, \quad v = t + r^* \quad (6.125)$$

thus for u and v constant we have radial null-curves. The coordinate transform $(t, r) \rightsquigarrow (v, u)$ gives us the following line-segment

$$ds^2 = - \left(1 - \frac{r_0}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2 \quad (6.126)$$

clearly we have removed the singularity at $r = r_0$. travel inward, toward the singularity.

Inside the Black Hole

In these coordinates outgoing radial null-curves are described by

$$\frac{dv}{dr} = \frac{2}{\left(1 - \frac{r_0}{r}\right)} \quad (6.127)$$

and infalling null-curves have $dv = 0$, we see that inside the black hole both infalling and outgoing null curves have negative $\frac{dv}{dr}$, which implies that light as well as massive objects are forced to travel towards the singularity if they are inside the event horizon. Interestingly, because there is a well-defined direction of the passage of r (always decreasing) we can think of r as a time-coordinate, whereas t becomes a spatial coordinate. This is also seen from the Schwarzschild metric, which is valid inside the black hole, as g_{tt} and g_{rr} have opposite signs inside and outside the black hole.

The singularity at the origin is a physical singularity in spacetime, therefore, we cannot remove it by changing coordinates. We can see this by finding an covariant quantity that describes the curvature of the spacetime, which diverges at the origin. We use the Kretschmann scalar for this:

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = \frac{12r_0^2}{r^6} \quad (6.128)$$

which clearly diverges at the origin.

Kruskal-Szekeres Coordinates

Our coordinate systems so far have not had global time coordinates; we needed to use different time coordinates inside and outside of the black hole. This is exactly what Kruskal-Szekeres Coordinates solve. They defined using the Eddington-Finkelstein quantities, u and v :

$$\mathcal{T} = \frac{1}{2} \left(e^{\frac{v}{2r_0}} - e^{-\frac{u}{2r_0}} \right), \quad \mathcal{R} = \frac{1}{2} \left(e^{\frac{v}{2r_0}} + e^{-\frac{u}{2r_0}} \right) \quad (6.129)$$

Using which we can write out the line-segment

$$ds^2 = \frac{4r_0^3}{r} e^{-\frac{r}{r_0}} \left(-d\mathcal{T}^2 + d\mathcal{R}^2 \right) + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (6.130)$$

where $r(\mathcal{I}, \mathcal{R})$ is defined through

$$\mathcal{I}^2 - \mathcal{R}^2 = \frac{r_0 - r}{r_0} e^{\frac{r}{r_0}} \quad (6.131)$$

There are four interesting parts of spacetime:

$$r = r_0 \quad \leftrightarrow \quad \mathcal{R} = \pm \mathcal{I} \quad (6.132)$$

$$r > r_0 \quad \leftrightarrow \quad |\mathcal{R}| > |\mathcal{I}| \quad (6.133)$$

$$r < r_0 \quad \leftrightarrow \quad |\mathcal{R}| < |\mathcal{I}| \quad (6.134)$$

$$r = 0 \quad \leftrightarrow \quad \mathcal{I}^2 = 1 + \mathcal{R}^2 \quad (6.135)$$

the final part is rather strange; the singularity is now given by two hyperbolae.

Gravitational Redshift

Once again we use

$$\frac{dt}{d\tau} = \frac{E}{1 - \frac{r_0}{r}} \quad (6.136)$$

putting this into the line-segment for a radially inward falling massive object:

$$\frac{dr}{d\tau} = \sqrt{E^2 - 1 + \frac{r_0}{r}} \quad (6.137)$$

The change in u and v are given by

$$du = \left(\frac{E + \sqrt{E^2 - 1 + \frac{r_0}{r}}}{1 - \frac{r_0}{r}} \right) d\tau, \quad dv = \left(\frac{E - \sqrt{E^2 - 1 + \frac{r_0}{r}}}{1 - \frac{r_0}{r}} \right) d\tau \quad (6.138)$$

which we obtain by simply using the above expressions for dt and dr .

The distant observer, that measures t , has constant $r = r_\infty$. Thus in his coordinate system $\Delta u = \Delta\tau_\infty = \Delta v$. However, r is not constant for the inward falling observer, however, we assume that it is approximately constant for the duration $\Delta\tau$. For an outgoing null-curve u is constant, therefore we can conclude that Δu must be the same for the observer at infinity and the inward falling observer, thus

$$\Delta\tau_\infty = \left(\frac{E + \sqrt{E^2 - 1 + \frac{r_0}{r}}}{1 - \frac{r_0}{r}} \right) \Delta\tau \quad (6.139)$$

Thus $\Delta\tau_\infty > \Delta\tau$; there is a gravitational redshift.

Similarly for radially inward falling null-curve v is constant, which means that the infalling observer experiences a change in period of light sent from the distant observer:

$$\Delta\tau_\infty = \left(\frac{E - \sqrt{E^2 - 1 + \frac{r_0}{r}}}{1 - \frac{r_0}{r}} \right) \Delta\tau \quad (6.140)$$

For $E > 1$ the factor on the right-hand side is always smaller than $\frac{1}{2}$, therefore for $E > 1$ we always have a gravitational redshift. However, for $E < \frac{1}{2}$ there is both blueshift and redshift, at the Schwarzschild radius there is always a blueshift for $E < \frac{1}{2}$.

Two fixed observers

If both the distant observer and the observer near the black hole have fixed positions we just get

$$d\tau^2 = \left(1 - \frac{r_0}{r}\right) d\tau_\infty^2 \quad (6.141)$$

Thus there is a redshift for light sent *to* the distant observer, and blueshift for light sent *from* the distant observer.

Observations of Black Holes

We differentiate between two groups of black holes; stellar mass black holes with masses of $3m_\odot$ to $100m_\odot$, and supermassive black holes with masses ranging from 10^6m_\odot to $10^{10}m_\odot$.

Stellar black holes are formed from collapsing remnants of supernovae. How supermassive black holes form is a bit of a mystery, though astrophysicists presume that they form from stellar black holes that collide and fuse, and slowly fall towards the centre of galaxies.

Black holes are difficult to observe directly, because they absorb all light. Before gravitational wave methods were developed we could only measure them indirectly; by looking at the trajectories of stars nearby.

6.8.2 *Kerr Black Hole*

The Schwarzschild metric is spherically symmetric, which means that it cannot rotate. Kerr black holes are allowed to rotate, so we no longer require a spherically symmetric metric, however, we still require that the metric is independent of time and that the metric is a solution to Einstein's Equations.

The line-element given by the Kerr metric is

$$\begin{aligned} ds^2 = & - \left(1 - \frac{r_0 r}{\Sigma}\right) dt^2 - \frac{2ar_0 r}{\Sigma} \sin^2 \theta dt d\phi \\ & + \frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2 \theta}{\Sigma} \sin^2 \theta d\phi^2 + \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2 \end{aligned} \quad (6.142)$$

where

$$\Sigma = r^2 + a^2 \cos^2 \theta, \quad \Delta = r^2 - r_0 r + a^2 \quad (6.143)$$

The event horizon is there where g_{rr} blows up, as was the case for the Schwarzschild black hole. For the Kerr Black hole this happens at

$$r_\pm = \frac{r_0}{2} \left(1 \pm \sqrt{1 - \frac{4a^2}{r_0^2}}\right) \quad (6.144)$$

The radius that is reached first by physical objects is r_+ , therefore we treat that as the event horizon. The geodesic equation for null curves us an expression for the effective (radial) speed of light:

$$\left(\frac{dr}{dt}\right)^2 = -\frac{g_{tt} + 2g_{t\phi}\frac{d\phi}{dt} + g_{\phi\phi}\left(\frac{d\phi}{dt}\right)^2 + g_{\theta\theta}\left(\frac{d\theta}{dt}\right)^2}{g_{rr}} \quad (6.145)$$

at the event horizon $g_{rr} \rightarrow \infty$, whereas the other metric-elements remain finite; this implies that light cannot escape the black hole, just like what we saw for the Schwarzschild black hole.

Ergoregion and frame dragging

The g_{tt} element of the metric is zero when $\Sigma = r_0 r$, which is satisfied for

$$r_{\text{ergo}} = \frac{r_0}{2} \left(1 \pm \sqrt{1 - \frac{4a^2}{r_0^2} \cos^2 \theta} \right) \quad (6.146)$$

the curve described by the $+$ solution is outside of the event horizon, therefore, it is accessible to our universe. Interestingly enough, inside the ergoregion we have

$$\begin{aligned} -2g_{t\phi}\frac{d\phi}{dt} &= \left(\frac{d\tau}{dt}\right)^2 + g_{tt} + g_{\phi\phi}\left(\frac{d\phi}{dt}\right)^2 \\ &\quad + g_{rr}\left(\frac{dr}{dt}\right)^2 + g_{\theta\theta}\left(\frac{d\theta}{dt}\right)^2 \end{aligned} \quad (6.147)$$

the right hand side is positive (not non-negative!) and $g_{t\phi} < 0$, therefore we have that $\dot{\phi} > 0$; there is a well defined direction in which everything inside the ergoregion must move, just like we saw *inside* the event horizon of the Schwarzschild black hole, however, the ergoregion is accessible to us, *and* it can be left! This effect is called frame dragging; the rotation of the Kerr black hole is forcing spacetime to corotate.

To further understand this, let us look at the parabola above. Specifically let us look at null curves that travel only in the ϕ direction, hence $dr = 0$ and $d\theta = 0$, and because they are null curves $d\tau = 0$, thus

$$-g_{\phi\phi}\Omega^2 - 2g_{t\phi}\Omega = g_{tt} \quad (6.148)$$

hence

$$\Omega_{\pm} = \frac{-g_{t\phi} \pm \sqrt{g_{t\phi}^2 - g_{\phi\phi}g_{tt}}}{g_{\phi\phi}} = \frac{-g_{t\phi} \pm \sin \theta \sqrt{\Delta}}{g_{\phi\phi}} \quad (6.149)$$

clearly massive objects must move at angular frequencies between the two extremes:

$$\Omega_- < \Omega < \Omega_+ \quad (6.150)$$

However as we decrease the distance from the centre toward r_+ the two frequencies will converge towards the same value, which we call the angular frequency of the black hole:

$$\Omega_H = \Omega_{\pm}(r_+, \theta) = \frac{a}{r_+ r_0} \quad (6.151)$$

6.8.3 *Asymptotically Flat metrics*

Both the Schwarzschild metric and the Kerr metric are asymptotically flat; at large distances from the black hole we can approximate the metric to be the Minkowski metric. This property is very important and makes intuitive sense, there may be a SMBH at the centre of the universe, but we do not expect it to prevent us from treating the space between the earth and the moon as approximately flat.

6.8.4 *Black Hole uniqueness*

The Schwarzschild and Kerr metric have the following properties:

- They are asymptotically flat
- There is a event horizon
- Outside the event horizon the metric is a solution to the vacuum Einstein Equations
- The metric is stationary

Amazingly there can exist no other metrics that satisfy these four conditions! This means that the two black holes we have discussed are the only two that satisfy these conditions and if there exist any other black holes in the universe it must be because at least one of these conditions is not satisfied. Thus M and J are the only two parametres we need to fully classify black holes.

6.8.5 *Cosmic Censorship Hypothesis*

Rotating black holes with $a > \frac{r_0}{2}$ would have a singularity that in principle is accessible to us; it is not shielded by a event horizon, which seems wrong. The *Cosmic Censorship Hypothesis* states that this is impossible, but does not have any grounds to base it on – it may very well be that these naked singularities exist in nature.

6.8.6 *Black Hole Mechanics*

Schwarzschild Black Hole

The surface gravity of a black hole, κ , is defined as the magnitude of the acceleration required for an object to be stationary at the Schwarzschild radius of a black hole, measured by the distant observer:

$$\sqrt{g_{\mu\nu}a_{(t)}^\mu a_{(t)}^\nu} = \frac{r_0}{2r^2} \quad (6.152)$$

hence

$$\kappa = \frac{1}{4GM} \quad (6.153)$$

The next constant we need to define is the area of the black hole, which is just

$$\mathcal{A} = 4\pi r_0^2 \quad (6.154)$$

using these quantities we can formulate the first law of black hole mechanics:

FIRST LAW OF BLACK HOLE MECHANICS (SCHWARZSCHILD)

Consider a small perturbation to the black hole, such that when it settles down again to a new stationary state it is again described as a Schwarzschild Black Hole. Then the change in mass M and area \mathcal{A} obey

$$\delta M = \frac{\kappa}{8\pi G} \delta \mathcal{A} \quad (6.155)$$

Kerr Black Hole

For a Kerr Black hole things become more complicated. The surface gravity becomes

$$\kappa = \frac{r_+ - r_-}{2r_0 r_+} \quad (6.156)$$

and the area we can get by r and θ in the line element, and then taking the limit $r \rightarrow r_+$

$$ds^2 = \frac{(r_+^2 + a^2)^2}{\Sigma} \sin^2 \theta d\phi^2 + \Sigma d\theta^2 \quad (6.157)$$

the Jacobian is the square root of determinant of the metric

$$\mathcal{A} = 2\pi(r_+^2 + a^2) \int_0^\pi d\theta \sin \theta = 4\pi(r_+^2 + a^2) \quad (6.158)$$

which can be rewritten to

$$\mathcal{A} = 8\pi G^2 M^2 \left(1 + \sqrt{1 - \frac{J^2}{G^2 M^4}} \right) \quad (6.159)$$

From which the first law for Kerr black holes follows

FIRST LAW OF BLACK HOLE MECHANICS (KERR)

Consider a small perturbation to a black hole. When it again settles down to a new stationary Kerr black hole the quantities M , J , Ω_H , κ and \mathcal{A} obey

$$dM = \frac{\kappa}{8\pi G} \delta \mathcal{A} + \Omega_H \delta J \quad (6.160)$$

where

$$\Omega_H \equiv \frac{a}{r_+ r_0} \quad (6.161)$$

Second Law of Black Hole Mechanics

The energy

$$E = -mg_{\mu\nu}T^\mu \frac{dx^\nu}{d\tau}, \quad T^\mu = \delta_t^\mu \quad (6.162)$$

is a conserved quantity far away from the black hole. However, in the ergoregion of a Kerr black hole, a part of this energy is spent on the forced rotation, due to frame dragging and the fact that t no longer is a good time coordinate it is not E that is conserved, but rather

$$\mathcal{E} = -mg_{\mu\nu}\chi^\mu \frac{dx^\nu}{d\tau}, \quad \chi^\mu = \delta_t^\mu + \Omega_H \delta_\phi^\mu \quad (6.163)$$

these two quantities are related through the angular momentum:

$$\mathcal{E} = E - \Omega_H L \quad L \equiv mg_{\phi\nu} \frac{dx^\nu}{d\tau} \quad (6.164)$$

The energy \mathcal{E} should be positive just before as it reaches the surface of the black hole, which using the conservation of energy gives us that

$$\delta M > \Omega_H \delta J \quad (6.165)$$

which implies that

SECOND LAW OF BLACK HOLE MECHANICS

The area A of the event horizon of a black hole can not decrease as a function of time (measured asymptotically) under any process that can be described by general relativity.

6.9 COSMOLOGY

6.9.1 Friedrich-Lemaître-Robertson-Walker metric

Let us assume there exist a universal time; *cosmic standard time*, t , which can parametrise the evolution of the universe. Let us use spherical coordinates for the spatial part:

$$x^\mu = (t, r, \theta, \phi) \quad (6.166)$$

Now we make two assumptions about the geometry of the entire universe:

- **HOMOGENEITY:** The geometry should look the same everywhere for a given t – invariant under spatial translations
- **ISOTROPY:** Given a point x^i , the universe should look the same in all directions; spherical symmetry about every point in the universe.

Naturally this does not hold on small scales; we require scales of at least order 100 million light years. These assumptions are what is referred to as the *cosmological principle*.

Due to the cosmological principle a line-element must be of the form

$$ds^2 = -dt^2 + g_{ij}dx^i dx^j \quad (6.167)$$

We see that stationary particles follow geodesics, provided that $\Gamma_{tt}^\mu = 0$. This is clearly satisfied because all terms that appear in the Christoffel Symbol are either zero or constant ($g_{t\mu} = -\delta_{t\mu}$). For fixed time the metric $g_{\mu\nu}|_{t=t_0}$ should describe homogenous, isotropic space. Let us consider only the spatial part of the metric

$$d\sigma^2 = \gamma_{ij}dx^i dx^j \quad (6.168)$$

Requiring that this is homogenous and isotropic means that it should be *maximally symmetric*. For such a space the spatial part of the Riemann Curvature Tensor takes the following form

$${}^{(\gamma)}R_{ijkl} = \frac{k}{a^2} (\gamma_{ik}\gamma_{jl} - \gamma_{il}\gamma_{jk}) \quad (6.169)$$

where $k \in \{-1, 0, 1\}$ is a dimensionless parametre that describes the type of space the universe is. a is a length parametre which is independent of x^i , but *can* be time-dependent. Furthermore we can show that the Ricci Tensor simplifies to

$${}^{(\gamma)}R_{ij} = \frac{2k}{a^2} \gamma_{ij} \quad (6.170)$$

using this we can solve the differential equations provided by the Christoffel Symbols and derive the *FLRW-metric*:

$$ds^2 = -dt^2 + a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right] \quad (6.171)$$

Let us change coordinates such that the singularity at $r = \frac{1}{k}$ disappears: $d\chi = \frac{dr}{\sqrt{1 - kr^2}}$ giving us

$$r(\chi) = \begin{cases} \sin \chi & \text{for } k = 1 \\ \chi & \text{for } k = 0 \\ \sinh \chi & \text{for } k = -1 \end{cases} \quad (6.172)$$

If $k = 1$ we say that the universe is *closed*: the metric describes a three-sphere. If $k = 0$ the universe is *flat* and finally if $k = -1$ we say the universe is *open*.

6.9.2 Hubble's Law

Depending on if $\dot{a} \equiv \frac{da}{dt}$ is positive, negative or zero, the FLRW-metric describes an expanding, contracting or static universe respectively.

Let our solar system be at the origin of the coordinate system. This would imply that every null-curve that reaches earth travels (roughly) on a radial null-curve. Hence

$$ds^2 = -dt^2 + a(t)^2 d\chi^2 \quad (6.173)$$

Integrating and using that $a(t)$ is approximately constant on times scales of the order of electromagnetic radiations periods;

$$\frac{T_f}{a(t_f)} = \frac{T_i}{a(t_i)} \quad (6.174)$$

where $T_f - T_i$ is the period of light. Defining the redshift parametre:

$$z \equiv \frac{\lambda_{\text{received}} - \lambda_{\text{emitted}}}{\lambda_{\text{emitted}}} \quad (6.175)$$

we can show that this can also be written as

$$z = \frac{a(t_f) - a(t_i)}{a(t_i)} \quad (6.176)$$

we see that if $z = 0$ the universe is static, if the universe expands $z > 0$ which implies that the light is *redshifted* and, finally, if $z < 0$ the light is *blueshifted*, corresponding to a contracting universe.

Consider a lightray emitted from a nearby galaxy; near enough that we can approximate the expansion of the universe as linear, then

$$a(t_i) \approx a(t_f) + \dot{a}(t_f)(t_i - t_f) \quad (6.177)$$

and so

$$z \approx \frac{\dot{a}(t_f)}{a(t_f)}(t_f - t_i) \quad (6.178)$$

in natural units $t_f - t_i = L$, the distance to the galaxy, thus there is a linear relation between the redshift parametre and the distance to nearby galaxies:

HUBBLE'S LAW

For nearby galaxies the redshift parametre, z , is proportional to the distance, L , to the galaxy to a good approximation:

$$z \approx H_0 L \quad (6.179)$$

where $H_0 \equiv \frac{\dot{a}(t_f)}{a(t_f)}$ is called *Hubble's constant*.

For nearby galaxies the time delay between emission and receival is so small (compared to cosmic time scales), that we can effectively say that the measured Hubble's constant is the current constant, and that it doesn't change in time.

The Hubble constant has been measured numerous times, however, with contradicting results. According to Planck 2015 $H_0 = (67 \pm 0.46) \text{ km s}^{-1} \text{ Mpc}^{-1}$, but more recent measurements $H_0 = (74 \pm 1.4) \text{ km s}^{-1} \text{ Mpc}^{-1}$. This contradicting empirical data is referred to as the *Hubble Tension*.

6.9.3 Friedmann Equations

Einstein's Equations for the description of the universe are rather ugly, because we can no longer assume that $T_{\mu\nu} = 0$; there matter, energy, pressure etc everywhere in the universe, which means the right-hand side of Einstein's Equations are non-zero. We've already discussed two types of fluids that can contribute to the energy-momentum tensor, which is of the form

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu + p\eta^{\mu\nu} \quad (6.180)$$

however, now we will add Dark Energy, and will motivate its existence shortly

$$\begin{aligned} \text{Dust:} & \quad p = 0 \\ \text{Photons:} & \quad p = \frac{1}{3}\rho \\ \text{Dark Energy:} & \quad p = -\rho \end{aligned}$$

When we derived Einstein's Field Equations we included a term which I chose to refer to as an integration constant:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu} + \Lambda g_{\mu\nu} \quad (6.181)$$

the Λ term can be thought of as an additional form of energy/momentum/pressure, which is there even if $T_{\mu\nu} = 0$. This is Dark Energy, and it contributes to the expansion of the universe:

$$\rho = -p = \frac{\Lambda}{8\pi G} \quad (6.182)$$

We can now solve Einstein's Field Equations for the expanding universe in the presence of Dark Energy, which results in the following

THE FRIEDMANN EQUATIONS

The scale factor $a(t)$ obeys the *Friedmann Equation*:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (6.183)$$

Additionally $a(t)$ obeys

$$\frac{\ddot{a}}{a} = -4\pi G \left(\rho + \frac{1}{3}p \right) \quad (6.184)$$

This is sometimes referred to as the second Friedmann Equation.

However, we can also look at the conservation of the energy momentum tensor:

$$0 = D_\mu T^\mu_t = - \left(\dot{\rho} + 3(\rho + p)\frac{\dot{a}}{a} \right) \quad (6.185)$$

ENERGY-MOMENTUM CONSERVATION

The energy momentum tensor is conserved if the energy density and pressures obey

$$\dot{\rho} = -\frac{3\dot{a}}{a}(\rho + p) \quad (6.186)$$

This equation holds for each type of cosmological fluid individually, and therefore also holds for the total universe. This is because in this model the fluids do no interact with each other.

It is noteworthy that the (first) Friedmann Equation together with energy-momentum conservation leads to the second Friedmann Equation.

6.9.4 Evolution of the Scale Factor

Let us introduce the time-dependent Hubble parametre $H(t) = \frac{\dot{a}}{a}$ as well as the critical density:

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G} \quad (6.187)$$

and the density parametre:

$$\Omega = \frac{8\pi G}{3H^2} \rho \quad (6.188)$$

with these the Friedmann Equation can be written

$$\Omega - 1 = \frac{k}{H^2 a^2} \quad (6.189)$$

thus we can associate the different values of Ω to the different models of the universe

$\Omega < 1$	$k = -1$	Open Universe
$\Omega = 1$	$k = 0$	Flat Universe
$\Omega > 1$	$k = 1$	Closed Universe

Composite Model

We will now consider a model of the universe, which consists of four different species of cosmological fluids

- Radiation, ρ_R . $w = \frac{1}{3}$
- Baryonic Matter, ρ_B . $w = 0$
- Dark Matter, ρ_{DM} . $w = 0$
- Dark Energy, ρ_Λ . $w = -1$

In this model the total energy-density is

$$\rho = \rho_R + \rho_B + \rho_{DM} + \rho_\Lambda \quad (6.190)$$

and

$$\Omega = \Omega_R + \Omega_B + \Omega_{DM} + \Omega_\Lambda \quad (6.191)$$

According to Planck 2015 the current values are

$\Omega_R^{(0)}$	< 0.001
$\Omega_B^{(0)}$	0.05
$\Omega_{DM}^{(0)}$	0.26
$\Omega_\Lambda^{(0)}$	0.69
Total	1.00

The total density of the universe is measured to be $\rho^{(0)} = 8.6 \times 10^{-26} \text{kg m}^{-3}$.

The Energy Momentum conservation condition for each species can be written as

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a} \quad (6.192)$$

which is solved if we let

$$\rho(t) \propto a^{-n} \quad (6.193)$$

where

$$n_R = 4, \quad n_B = n_{DM} = 3, \quad n_\Lambda = 0 \quad (6.194)$$

Note that $H = \frac{8\pi G}{3} \sum_i \rho_i > 0$, which implies that the universe is expanding. Additionally

$$\dot{H} = -4\pi G \sum_i (1+w_i)\rho_i \leq 0 \quad (6.195)$$

which implies that the Hubble parametre decreases as a function of time.

The universe is dominated more and more by dark energy, thus we expect that in the distant future we will have

$$a(t) = a_\Lambda e^{H_\Lambda t} \quad (6.196)$$

Past of the Universe

As we saw, $a(t)$ is related to the redshift parametre:

$$\frac{a(t_0)}{a(t)} = 1 + z \quad (6.197)$$

where t_0 is the time now and t is the time which we looking back at. Thus

$$\begin{aligned} \rho_R &= \Omega_R^{(0)}(1+z)^4, & \rho_M &= \Omega_M^{(0)}(1+z)^3, \\ \rho_\Lambda &= \Omega_\Lambda^{(0)}(1+z) \end{aligned} \quad (6.198)$$

where M now includes baryonic and dark matter. For a universe dominated by an energy density we have that $\rho \propto a^{-n}$, which implies that $a(t) \propto (t - t_*)^{\frac{2}{n}}$. This implies that there was a beginning of the universe. However, this is not the case for Λ -dominated universes, as their beginning lies at $t = -\infty$.

The dark energy dominated phase of the universe (now) lies between

$$\Lambda\text{--dominated} \quad 0 < z < 0.4$$

Whereas the matter dominated phase was at

$$\text{Matter -- dominated} \quad 0.4 < z < 3600$$

and finally

$$\text{Radiation -- dominated} \quad 3600 < z$$

Temperature of the Universe

Currently the temperature of the Universe is $T_0 = 2.726\text{K}$. As we look back in time, this temperature will increase with z , proportionally to $(1 + z)$, as the temperature is proportional to the wavelength of the cosmic background radiation:

$$T(z) = T_0(1 + z) \quad (6.199)$$

6.10 LINEARISED GRAVITY

Consider the limit where $g_{\mu\nu}$ differs slightly from the Minkowski metric:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| \ll 1 \quad (6.200)$$

plugging this into the Christoffel Symbol and only keeping terms of linear order or less:

$$\Gamma_{\mu\nu}^\rho = \frac{1}{2} (\partial_\mu h_\nu^\rho + \partial_\nu h_\mu^\rho - \partial^\rho h_{\mu\nu}) \quad (6.201)$$

The geodesic equation for massive objects:

$$\frac{d^2 x^\rho}{d\tau^2} = - \left(\partial_\mu h_\nu^\rho - \frac{1}{2} \partial^\rho h_{\mu\nu} \right) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \quad (6.202)$$

The Riemann Curvature Tensor becomes

$$R_{\mu\nu} = \frac{1}{2} (\partial_\mu \partial_\rho h_\nu^\rho + \partial_\nu \partial_\rho h_\mu^\rho - \square^2 h_{\mu\nu} - \partial_\mu \partial_\nu h_\rho^\rho) \quad (6.203)$$

where $\square^2 = \partial_\mu \partial^\mu$. This expression for the curvature tensor only includes terms of linear order in h . We can freely Gauge transform our metric without it affecting the curvature tensor (though it might look different, we just transformed it using the rules for a $(0, 2)$ -tensor). In a suitable gauge the Riemann Curvature Tensor just becomes

$$R_{\mu\nu} = -\frac{1}{2} \square^2 h_{\mu\nu} \quad (6.204)$$

in this gauge we have that

$$\partial^\mu \left(h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h_\rho^\rho \right) = 0 \quad (6.205)$$

(up to linear order in h).

LINEARISED EINSTEIN FIELD EQUATIONS:

In the weak field limit of gravity Einsteins Field Equations can be written as

$$\square^2 h_{\mu\nu} = -16\pi G \left(T_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} \eta^{\rho\sigma} T_{\rho\sigma} \right) \quad (6.206)$$

up to linear order in $h_{\mu\nu}$. We impose the Lorentz-Gauge on $h_{\mu\nu}$, which implies that

$$\partial^\mu \left(h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h_\rho^\rho \right) = 0 \quad (6.207)$$

6.10.1 Gravitational Waves

In vacuum $T_{\mu\nu} = 0$ which implies that the metric must satisfy the D'Alembert equation:

$$\square^2 h_{\mu\nu} = 0 \quad (6.208)$$

We know the form the solution to this:

$$h_{\mu\nu} = A_{\mu\nu} \exp(ik_\rho x^\rho) \quad (6.209)$$

the Lorentz Gauge implies that

$$k^\mu A_{\mu\nu} = \frac{1}{2} k_\nu \eta^{\rho\sigma} A_{\rho\sigma} \quad (6.210)$$

Additionally the D'Alembert equation tells us that

$$k_\mu k^\mu = 0 \quad (6.211)$$

Consider a wave going in the x^3 direction, for it we must have

$$k^\mu = (\omega, 0, 0, \omega) \quad (6.212)$$

which in turn implies that

$$A_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{11} & A_{12} & 0 \\ 0 & A_{12} & -A_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (6.213)$$

clearly there are two polarisations, both of which are orthogonal to the direction of propagation – just like electromagnetic radiation. Setting

$$h_{11}(t) = B_1 \cos(\omega t + \chi_1), \quad h_{12}(t) = B_2 \cos(\omega t + \chi_2) \quad (6.214)$$

we can calculate the proper distance between two test particles whose position at $t = 0$ are known to be

$$(x_A, y_A), \quad (x_A + L_0 \cos \theta, y_A + L_0 \sin \theta) \quad (6.215)$$

RELATIVE MOTION OF TEST PARTICLES

Consider two test particles that lie in the plane perpendicular to the propagation of the gravitational wave. Their position in this plane are given by Equation 6.215. The proper distance between these particles is given by

$$\frac{L(t)}{L_0} = \left[1 + \frac{1}{2} h_{11}(t) \cos(2\theta) + \frac{1}{2} h_{12}(t) \sin(2\theta) \right] \quad (6.216)$$

where

$$h_{11}(t) = B_1 \cos(\omega t + \chi_1) \quad (6.217)$$

$$h_{12}(t) = B_2 \cos(\omega t + \chi_2) \quad (6.218)$$

Using interferometry one can measure this effect, however, only for gravitational waves whose wavelengths are much greater than the size of the device used. The issue is that we need very large devices in order to obtain the measurement accuracy required to make these measurements, therefore, the size and energy of gravitational events that can be measured is limited.

ADVANCED QUANTUM MECHANICS

7.1 FUNDAMENTAL CONCEPTS

7.1.1 *Stern-Gerlach Experiment*

Quantum mechanical particles have intrinsic magnetic moments due to their spin angular momentum:

$$\boldsymbol{\mu} \propto \mathbf{S} \quad (7.1)$$

Devices such as the *Stern-Gerlach Apparatus* can measure the z-component of the magnetic moment and hence spin. Subsequent measurements with Stern-Gerlach Apparatuses with non-commuting spins have results that are quite different from what we'd expect with our classical intuition.

7.1.2 *Bras, Kets and Operators*

Quantum mechanical states are described by vectors in an infinite complex vector space (Hilbert Space). The vectors, kets, in this space are denoted by $|\alpha\rangle$ and their dual vectors, bras, are denoted by $\langle\alpha|$. These kets describe an abstract vector, whose time evolution are governed by the Schrödinger Equation

$$i\hbar\partial_t|\alpha\rangle = \hat{H}|\alpha\rangle \quad (7.2)$$

This can also be written for the bras:

$$-i\hbar\partial_t\langle\alpha| = \hat{H}\langle\alpha| \quad (7.3)$$

where I've used that \hat{H} is a *hermitian operator*

The *inner product* of a bra and a ket gives you a scalar:

$$\langle\alpha|\beta\rangle = c \in \mathbb{C} \quad (7.4)$$

which tells you about the projection of β onto α .

The *outer product* gives you an operator:

$$|\alpha\rangle\langle\beta| \quad (7.5)$$

which is useful for changing basis

Hermitian Operators

A hermitian operator, \hat{A} , fulfills:

$$\hat{A}^\dagger = \hat{A} \quad (7.6)$$

The Hermiticity of an operator has two very useful implications:

- Real eigenvalues: hence observables are represented by hermitian operators
- Diagonalisability: There exists a basis in which \hat{A} is diagonal.

Projection Operators

Define the projection operator:

$$\Lambda_a = |a\rangle\langle a| \quad (7.7)$$

If $\{|a\rangle\}$ forms an orthonormal set, we have

$$\mathbb{1} = \sum \Lambda_a \quad (7.8)$$

and if $\hat{A}|a\rangle = a|a\rangle$, for $a \in \mathbb{R}$, we can write the operator in terms of its spectral decomposition:

$$\hat{A} = \sum_n a_n \Lambda_{a_n} \quad (7.9)$$

Due to the fact that $\{|a\rangle\}$ is an orthonormal basis we have that $\Lambda_a \Lambda_b = \delta_{a,b} \Lambda_a$ and hence given an $f(x)$ that has a power expansion we can define $f(\hat{A}) = \sum_m c_m \hat{A}^m$ which in turn gives us

$$f(\hat{A}) = \sum_n f(a_n) \Lambda_{a_n} = \sum_n f(a_n) |a_n\rangle\langle a_n| \quad (7.10)$$

Commutating Operators

Given two operators \hat{A} and \hat{B} , these operators commute given that

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = 0 \quad (7.11)$$

in which case there exists a basis $\{|a, b\rangle\}$ that is a simultaneous eigenbasis for both \hat{A} and \hat{B} :

$$\hat{A}|a, b\rangle = a|a, b\rangle \quad \hat{B}|a, b\rangle = b|a, b\rangle \quad (7.12)$$

if $[\hat{A}, \hat{B}] = 0$ then $[\hat{A}, f(\hat{B})] = 0$, which can be shown using the power expansion of $f(x)$.

Non-commuting operators

For non-commuting operators $[\hat{A}, \hat{B}] \neq 0$, which implies that there does not exist a simultaneous set of eigenkets. Additionally we can show that

$$[\hat{A}, f(\hat{B})] = [\hat{A}, \hat{B}] f'(\hat{B}) \quad (7.13)$$

Measurement

Consider a general state

$$|\psi\rangle = \sum_n c_n |n\rangle \quad (7.14)$$

where $\hat{A}|n\rangle = a_n|n\rangle$, $a_n \in \mathbb{R}$. If we prepare a state $|\psi\rangle$ and then measure \hat{A} , we will obtain one of the eigenvalues of \hat{A} , a_n , with probability:

$$P(a_n) = |\langle a_n|\psi\rangle|^2 \quad (7.15)$$

If we measure a_n we know for certain that $|\psi\rangle = |a_n\rangle$; we say the state *collapses* into $|a_n\rangle$. In the case that we have a degenerate eigenvalue we only know that a_n is measured, which implies that $|\psi\rangle$ collapses into the degenerate space, hence after the measurement $|\psi\rangle$ can be written as a linear combination of the bases that spans the degenerate space.

Change of Basis

Given two orthonormal bases, $\{|a_n\rangle\}$ and $\{|b_n\rangle\}$ of the Hilbert space in question, there exists a unitary transformation, \hat{U} such that

$$|b_n\rangle = \hat{U}|a_n\rangle \quad (7.16)$$

\hat{U} can be written in terms of the outer products defined above:

$$\hat{U} = \sum_n |b_n\rangle\langle a_n| \quad (7.17)$$

Matrix Representations of Operators

Operators can be expressed as matrices given a basis:

$$A_{ij} = \langle i|\hat{A}|j\rangle \quad (7.18)$$

clearly, A , is a diagonal matrix if $\{|i\rangle\}$ form its eigenbasis. Using the change of basis from above we can write similarity transformations as

$$A' = U^\dagger A U \quad (7.19)$$

In this way we can diagonalise matrices, if we choose the new basis to be the eigenbasis of \hat{A} .

Continuous Spectra

Some operators have continuous eigenvalues and we have to treat these differently from discrete operators. In general we replace sums by integrals and Kronecker deltas by Dirac deltas:

$$\mathbb{1} = \int d\alpha |\alpha\rangle\langle\alpha| \quad (7.20)$$

$$\langle\alpha|\beta\rangle = \delta(\alpha - \beta) \quad (7.21)$$

hence a general state can be written as

$$|\psi\rangle = \int d\alpha \langle\alpha|\psi\rangle |\alpha\rangle \quad (7.22)$$

Translation

The (infinitesimal) translation operator $\mathcal{J}(\mathbf{dx})$ takes an eigenket of the position operator and displaces it by \mathbf{dx} :

$$\mathcal{J}(\mathbf{dx})|\mathbf{x}\rangle = |\mathbf{x} + \mathbf{dx}\rangle \quad (7.23)$$

In order for this to preserve probabilities it must be unitary, which has the implication that

$$\mathcal{J}^\dagger(\mathbf{dx}) = \mathcal{J}(-\mathbf{dx}), \quad \mathcal{J}^{-1}(\mathbf{dx}) = \mathcal{J}(-\mathbf{dx}) \quad (7.24)$$

for infinitesimal displacements we can write out the first order approximation for the displacement operator:

$$\mathcal{J}(\mathbf{dx}) = \mathbb{1} - i\mathbf{K} \cdot \mathbf{dx} \quad (7.25)$$

where \mathbf{K} is hermitian due to the unitarity of \mathcal{J} . This hermitian operator is, in fact, the momentum operator: the momentum operator is the generator for displacement, as we would expect from classical mechanics:

$$\mathcal{J}(\mathbf{dx}) = \mathbb{1} - \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{dx} \quad (7.26)$$

by repeated applying this for finite displacements:

$$\mathcal{J}(\mathbf{a}) = \lim_{N \rightarrow \infty} \left(\mathbb{1} - \frac{i\mathbf{p} \cdot \mathbf{a}}{N\hbar} \right)^N \equiv \exp \left(-\frac{i\mathbf{p} \cdot \mathbf{a}}{\hbar} \right) \quad (7.27)$$

Due to the fact that the displacement operator is only a function of \mathbf{p} we see that $[\mathbf{p}, \mathcal{J}(\mathbf{a})] = 0$. Indeed:

$$\mathcal{J}(\mathbf{a})|\mathbf{p}\rangle = \exp \left(-\frac{i\hat{\mathbf{p}} \cdot \mathbf{a}}{\hbar} \right) |\mathbf{p}\rangle = \exp \left(-\frac{i\mathbf{p} \cdot \mathbf{a}}{\hbar} \right) |\mathbf{p}\rangle \quad (7.28)$$

7.1.3 *Wave Functions*

We can represent our abstract $|\alpha\rangle$ in the position and momentum:

$$\psi_\alpha(\mathbf{x}) \equiv \langle \mathbf{x} | \alpha \rangle, \quad \phi_\alpha(\mathbf{p}) \equiv \langle \mathbf{p} | \alpha \rangle \quad (7.29)$$

with these wavefunctions we can calculating expectation values, as long as we know how the operators operate on the wavefunctions. For instance, we can establish what the momentum operator looks like in position basis up to linear order in the infinitesimal displacement \mathbf{dx} :

$$\left(\mathbb{1} - \frac{i\mathbf{p} \cdot \mathbf{dx}}{\hbar} \right) |\alpha\rangle = \int d\mathbf{x}' |\mathbf{x}'\rangle \langle \mathbf{x}' | \mathcal{J}(\mathbf{dx}) | \alpha \rangle \quad (7.30)$$

$$= \int d\mathbf{x}' |\mathbf{x}'\rangle \langle \mathbf{x}' - \mathbf{dx} | \alpha \rangle \quad (7.31)$$

$$= \int d\mathbf{x}' |\mathbf{x}'\rangle \left(\langle \mathbf{x}' | \alpha \rangle - \mathbf{dx} \cdot \frac{\partial}{\partial \mathbf{x}'} \langle \mathbf{x}' | \alpha \rangle \right) \quad (7.32)$$

comparing the linear terms in \mathbf{dx} we get

$$\langle \mathbf{x}' | \mathbf{p} | \alpha \rangle = -i\hbar \nabla' \langle \mathbf{x}' | \alpha \rangle \quad (7.33)$$

where we've used $\langle \mathbf{x} | \mathbf{x}' \rangle = \delta(\mathbf{x}' - \mathbf{x})$.

Fourier Transform

The momentum- and position-space wavefunctions are related through the Fourier-transform:

$$\phi_\alpha(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int d^3\mathbf{x} \exp\left(-\frac{i\mathbf{p} \cdot \mathbf{x}}{\hbar}\right) \psi_\alpha(\mathbf{x}) \quad (7.34)$$

7.2 QUANTUM DYNAMICS

Let us assume we know the ket of some state at $t = t_0$: $|\alpha, t_0\rangle$, then there must exist a unitary operator that transforms $|\alpha, t_0\rangle \rightarrow |\alpha, t_0; t\rangle$, that is, the stateket at $t = t$. This unitary operator is the time-evolution operator:

$$|\alpha, t_0; t\rangle = \mathcal{U}(t, t_0)|\alpha, t_0\rangle \quad (7.35)$$

For infinitesimal time differences $t \rightarrow t + dt$:

$$|\alpha, t; t + dt\rangle = \mathcal{U}(t + dt, t)|\alpha, t\rangle \quad (7.36)$$

this gives us

$$\left(\mathbb{1} + dt \frac{\partial}{\partial t}\right) |\alpha, t\rangle = (\mathbb{1} - i\hat{\Omega}dt) |\alpha, t\rangle \quad (7.37)$$

hence

$$i \frac{\partial}{\partial t} |\alpha, t\rangle = \hat{\Omega} |\alpha, t\rangle \quad (7.38)$$

the operator $\hat{\Omega} = \frac{1}{\hbar} \hat{H}$ where \hat{H} is the Hamiltonian operator, hence

$$i\hbar \frac{\partial}{\partial t} |\alpha, t\rangle = \hat{H} |\alpha, t\rangle \quad (7.39)$$

we can use the same trick as we did for the displacement operator to derive a closed form expression for the time-evolution operator:

$$\mathcal{U}(t, t_0) = \exp\left(-\frac{i\hat{H}(t - t_0)}{\hbar}\right) \quad (7.40)$$

however, *this assumes that \hat{H} is not a function of time*. Generally if $\hat{H} = \hat{H}(t)$ and $[\hat{H}(t), \hat{H}(t')] \neq 0$ we can use the *Dyson Series*:

$$\mathcal{U}(t, t_0) = 1 + \sum_{N=1}^{\infty} \left(\frac{-i}{\hbar}\right)^N \int_{t_0}^t \left(\prod_{n=1}^{N-1} dt_n \int_{t_0}^{t_n}\right) dt_N \prod_{n=1}^N \hat{H}(t_n) \quad (7.41)$$

which simplifies if $[\hat{H}(t), \hat{H}(t')] = 0 \quad \forall t, t'$ to

$$\mathcal{U}(t, t_0) = \exp\left(-\frac{i}{\hbar} \int_{t_0}^t dt \hat{H}(t)\right) \quad (7.42)$$

If we decompose $|\alpha\rangle$ into energy eigenstates the time evolution operator can easily be evaluated using the power-expansion method (we assume again that H is independent of time):

$$|\alpha, t_0\rangle = \sum_n c_n |a_n\rangle \quad (7.43)$$

then

$$|\alpha, t_0; t\rangle = \sum_n c_n \mathcal{U}(t, t_0) |a_n\rangle = \sum_n c_n \exp\left(-\frac{iE_n t}{\hbar}\right) |a_n\rangle \quad (7.44)$$

for $\hat{H}|a_n\rangle = E_n|a_n\rangle$

Hamiltonian in position space

The Hamiltonian operator in position space takes the form

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x}) \quad (7.45)$$

which can be derived from the classical Hamiltonian function by using that $\mathbf{p} \rightsquigarrow -i\hbar\nabla$

Note that the potential term only looks like this for local potentials, *i.e.* ones for which

$$\langle \mathbf{x}'' | V(\mathbf{x}) | \mathbf{x}' \rangle = V(\mathbf{x}') \delta(\mathbf{x}' - \mathbf{x}'') \quad (7.46)$$

7.2.1 Heisenberg Picture

In the Schrödinger picture operators have no Hamiltonian evolution, whereas the states do. In the Heisenberg picture states are constant whereas the operators evolve due to the Hamiltonian operator, we can understand this by looking at a time-dependent expectation value:

$$\langle A \rangle(t) = \langle \alpha | \mathcal{U}^\dagger(t, t_0) \hat{A} \mathcal{U}(t, t_0) | \alpha \rangle \quad (7.47)$$

the Schrödinger picture lets the time-evolution operate on the states, but the Heisenberg operator instead defines

$$A^{(H)}(t) = \mathcal{U}^\dagger(t, t_0) \hat{A}^{(S)} \mathcal{U}(t, t_0) \quad (7.48)$$

and leaves $|\alpha\rangle$ unchanged (they do not evolve in time). In the Heisenberg picture the time-evolution of operators is governed by the *Heisenberg equation of motion*:

$$\frac{dA^{(H)}}{dt} = \frac{1}{i\hbar} [A^{(H)}, H] \quad (7.49)$$

which is almost identical to the time evolution of functions in classical physics, if we relate the commutators to Poisson brackets.

Transition Amplitude

The time-dependent transition amplitude from $|a\rangle$ to $|b\rangle$ is given by the inner product between the time-dependent $|a(t)\rangle = \mathcal{U}|a\rangle$ and $|b\rangle$:

$$\langle b | \mathcal{U} | a \rangle \quad (7.50)$$

7.2.2 Harmonic Oscillator

The Hamiltonian for the harmonic oscillator is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \quad (7.51)$$

which can be rewritten as

$$\hat{H} = \frac{\hbar\omega}{2} (\hat{P}^2 + \hat{X}^2) \quad (7.52)$$

where $\hat{P} = \frac{\hat{p}}{\sqrt{m\hbar\omega}}$ and $\hat{X} = \sqrt{\frac{m\omega}{\hbar}}\hat{x}$. We have that

$$[\hat{P}, \hat{X}] = i \quad (7.53)$$

hence we can define

$$a = \frac{1}{\sqrt{2}} (\hat{X} + i\hat{P}), \quad a^\dagger = \frac{1}{\sqrt{2}} (\hat{X} - i\hat{P}) \quad (7.54)$$

which which we can write

$$\hat{H} = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \quad (7.55)$$

Defining the number operator, $\hat{N} = a^\dagger a$, we see that because $[\hat{H}, \hat{N}] = 0$ we can denote the eigenstates of \hat{H} by the quantum number n , which tells us about the number of energy quanta there are in stored in the oscillator. We can show that

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad a |n\rangle = \sqrt{n} |n-1\rangle \quad (7.56)$$

These operators are very powerful, they can, for instance be used to find the wavefunctions. For example the wavefunction $\psi_0(x) = \langle x|0\rangle$ must satisfy

$$\left(x + x_0^2 \frac{d}{dx} \right) \psi_0(x) = 0, \quad x_0 \equiv \sqrt{\frac{\hbar}{m\omega}} \quad (7.57)$$

7.2.3 Interpretation of the wavefunction

The standard interpretation of the Schrödinger wavefunction is that its modulus squared is a probability density:

$$\rho(\mathbf{x}, t) = |\psi(\mathbf{x}, t)|^2 \quad (7.58)$$

which indeed leads to a continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \quad (7.59)$$

where

$$\mathbf{j} = - \left(\frac{i\hbar}{2m} \right) [\psi^* \nabla \psi - (\nabla \psi^*) \psi] \quad (7.60)$$

Consider now the wavefunction as a real function times a phase:

$$\psi(\mathbf{x}, t) = \sqrt{\rho(\mathbf{x}, t)} \exp \left(\frac{iS(\mathbf{x}, t)}{\hbar} \right) \quad (7.61)$$

this tells us that

$$\mathbf{j} = \frac{\rho \nabla S}{\hbar} \quad (7.62)$$

So the probability flux is due to variations in the phase of ψ . Additionally, writing the Schrödinger Equation out for this function we get, for $\sqrt{\rho} = f$

$$-\left(\frac{\hbar^2}{2m}\right)\left(\nabla^2 f + \frac{2i}{\hbar}\nabla f \cdot \nabla S - \frac{1}{\hbar^2}f|\nabla S|^2\right) \quad (7.63)$$

$$-\frac{i\hbar}{2m}f\nabla^2 S + fV = i\hbar\left(\frac{\partial f}{\partial t} + \left(\frac{i}{\hbar}\right)f\frac{\partial S}{\partial t}\right) \quad (7.64)$$

Let us assume the curvature of S is much smaller than the gradient:

$$\hbar|\nabla^2 S| \ll |\nabla S|^2 \quad (7.65)$$

in this limit the Schrödinger Equation simplifies to

$$H(\mathbf{x}, \nabla S) + \frac{\partial S}{\partial t} = 0 \rightsquigarrow \frac{1}{2m}|\nabla S|^2 + V(\mathbf{x}) + \frac{\partial S}{\partial t} = 0 \quad (7.66)$$

which is the Hamilton-Jacobi Equation

7.2.4 Propagators

Propagators are functions that relate the wavefunction at $(x^\mu)'$ to the wave function at x^μ . Let us consider a time-independent Hamiltonian, so the time-evolution operator has a simple form for eigenkets:

$$|\alpha, t'; t\rangle = \exp\left(-\frac{iH(t-t_0)}{\hbar}\right)|\alpha, t'\rangle \quad (7.67)$$

$$= \sum_a |a\rangle \langle a|\alpha\rangle \exp\left(-\frac{iE_a(t-t')}{\hbar}\right) \quad (7.68)$$

Multiplying by an identity $\mathbb{1} = \int d\mathbf{x}' |\mathbf{x}'\rangle \langle \mathbf{x}'|$ we get

$$\psi(\mathbf{x}, t) = \int d^3\mathbf{x}' \sum_a \langle \mathbf{x}|a\rangle \langle a|\mathbf{x}'\rangle \psi(\mathbf{x}', t') e^{-\frac{iE_a(t-t')}{\hbar}} \quad (7.69)$$

This relates the wavefunction at $\psi(\mathbf{x}', t')$ to the wavefunction at $\psi(\mathbf{x}, t)$. The quantity that propagates from (\mathbf{x}', t') to (\mathbf{x}, t) is the *propagator*:

$$K(\mathbf{x}, t; \mathbf{x}', t') \equiv \sum_a \langle \mathbf{x}|a\rangle \langle a|\mathbf{x}'\rangle e^{-\frac{iE_a(t-t')}{\hbar}} \quad (7.70)$$

more generally

$$K(\mathbf{x}, t; \mathbf{x}', t') \equiv \sum_a \langle \mathbf{x}|a\rangle \langle a|\mathbf{x}'\rangle \langle a|\mathcal{U}(t, t')|a\rangle \quad (7.71)$$

Does this still work in the adiabatic limit?

but this only holds if $[H(t), H(t')]$, because otherwise the eigenbasis $\{|a\rangle\}$ is time-dependent.

Propagators can be written as the inner product between the state at (\mathbf{x}', t') and (\mathbf{x}, t) :

$$K(\mathbf{x}, t; \mathbf{x}', t') = \langle \mathbf{x}, t | \mathbf{x}', t' \rangle \quad (7.72)$$

We can compose propagators:

$$K(\mathbf{x}, t; \mathbf{x}'', t'') = \int d\mathbf{x}' \langle \mathbf{x}, t | \mathbf{x}', t' \rangle \langle \mathbf{x}', t' | \mathbf{x}'', t'' \rangle \quad (7.73)$$

$t > t' > t''$

The composition property can be taken to the extreme, such that

$$\langle x_N, t_N | x_1, t_1 \rangle = \int dx_{N-1} \int dx_{N-2} \cdots \int dx_2 \quad (7.74)$$

$$\times \langle x_N, t_N | x_{N-1}, t_{N-1} \rangle \langle x_{N-1}, t_{N-1} | x_{N-2}, t_{N-2} \rangle \cdots \times \langle x_2, t_2 | x_1, t_1 \rangle \quad (7.75)$$

where $t_m - t_{m-1}$ is infinitesimal. This form integrates over every possible path that connects (x_1, t_1) to (x_N, t_N) .

Feynman's Formulation

The propagator defined above corresponds:

$$\langle x_2, t_2 | x_1, t_1 \rangle \sim \exp \left(i \int_{t_1}^{t_2} dt \frac{L_{\text{classical}}(x, \dot{x})}{\hbar} \right) \quad (7.76)$$

For compactness let

$$S_{n,n-1} \equiv \int_{t_{n-1}}^{t_n} dt L_{\text{classical}}(x, \dot{x}) \quad (7.77)$$

which means the total $\exp(iS_{N,1}/\hbar)$ can be written as

$$\exp \left(\frac{iS_{N,1}}{\hbar} \right) = \exp \left(\left(\frac{i}{\hbar} \right) \sum_{n=2}^N S_{N,n-1} \right) \quad (7.78)$$

However to calculate $\exp \left(\frac{iS_{N,1}}{\hbar} \right)$ we need to have a parametrisation for the path taken, therefore we now also need to sum over all possible paths:

$$\langle x_N, t_N | x_1, t_1 \rangle \sim \sum_{\text{all paths}} \exp \left(\frac{iS_{N,1}}{\hbar} \right) \quad (7.79)$$

The full form of this is **Feynman's Path Integral**:

$$\langle x_N, t_N | x_1, t_1 \rangle = \int_{x_1}^{x_N} \mathcal{D}[x(t)] \exp \left[\int_{t_1}^{t_N} dt \frac{L_{\text{classical}}}{\hbar} \right] \quad (7.80)$$

where

$$\mathcal{D}[x(t)] = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi i \hbar \Delta t} \right)^{\left(\frac{N-1}{2}\right)} \prod_{n=N}^2 \int dx_n \quad (7.81)$$

7.2.5 Gauge Transformations in Electromagnetism

Using the scalar and vector potentials from electromagnetism we can formulate the Hamiltonian of charged particles in electromagnetic fields:

$$H = \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A} \right)^2 + e\phi \quad (7.82)$$

however, we need to be careful when we multiply the product out, because in general the momentum and vector potential operators do not commute:

$$\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 = |\mathbf{p}|^2 + \left(\frac{e}{c}\right)^2 |\mathbf{A}|^2 - \frac{e}{c}(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) \quad (7.83)$$

The canonical momentum, \mathbf{p} , is no longer the conjugate variable to \mathbf{x} , however, we can define the kinematic momentum:

$$\mathbf{\Pi} \equiv \mathbf{p} - \frac{e\mathbf{A}}{c} \quad (7.84)$$

Note that it is the kinematic momentum and not the canonical momentum that is a conserved quantity. The continuity equation still looks the same, if we redefine the probability flux:

$$\mathbf{j} = -\left(\frac{i\hbar}{2m}\right)[\psi^*\nabla\psi - (\nabla\psi^*)\psi] - \left(\frac{e}{mc}\right)\mathbf{A}\rho \quad (7.85)$$

Maxwell's equations are invariant under Gauge transformations:

$$\phi \rightsquigarrow \phi - \frac{1}{c}\frac{\partial\Lambda}{\partial t}, \quad \mathbf{A} \rightsquigarrow \mathbf{A} + \nabla\Lambda \quad (7.86)$$

The equations of motion in quantum mechanics are invariant under Gauge transformations too, however, these transformations to leave a phase in the state:

$$|\alpha\rangle \rightsquigarrow \exp\left(\frac{ie\Lambda}{\hbar c}\right)|\alpha\rangle \quad (7.87)$$

7.2.6 Aharonov-Bohm Effect

The fact that the Schrödinger Equation depends on the scalar and vector potentials and not the electromagnetic fields implies that there are instances in where a wavefunction can be influenced by the electromagnetic even though both $\mathbf{E} = 0$ and $\mathbf{B} = 0$. Consider for example a region where $B \neq 0$, however, this region is shielded such that the wavefunction cannot enter it. The vector potential is nonzero outside of the region. Two paths that fly past the impenetrable region will obtain a different phase depending on whether they pass the region in the same way (with the region on their right or left). This phase difference results in a peculiar interference pattern, which can be and has been measured.

7.3 THEORY OF ANGULAR MOMENTUM

We have considered the displacement-transformation previously, which we described with a unitary operator and in looking at infinitesimal displacements we found the generator for displacements, the momentum operator. We will now do the same for rotations. Consider an infinitesimal rotation, $d\phi$, about the unit vector $\hat{\mathbf{n}}$. The rotation operator can be expanded to first order in the generator, which we will call \mathbf{J} :

$$\mathcal{D}(\hat{\mathbf{n}}, d\phi) \approx \mathbb{1} - i\left(\frac{\mathbf{J} \cdot \hat{\mathbf{n}}}{\hbar}\right)d\phi \quad (7.88)$$

we can once again to N infinitesimal subsequent rotations and let $N \rightsquigarrow \infty$ to get a finite rotation:

$$\mathcal{D}(\hat{\mathbf{n}}, \phi) = \exp \left(-i \left(\frac{\mathbf{J} \cdot \hat{\mathbf{n}}}{\hbar} \right) \phi \right) \quad (7.89)$$

Due to the fact that we know how physical rotations commute, we can require that \mathcal{D} satisfies the same commutation relations, which tells us how the \mathbf{J} operators commute. For instance for rotations by an infinitesimal angle, ε , we have that

$$R_x(\varepsilon)R_y(\varepsilon) - R_y(\varepsilon)R_x(\varepsilon) = R_z(\varepsilon^2) - 1 + O(\varepsilon^4) \quad (7.90)$$

Hence we can require that

$$\mathcal{D}(\hat{\mathbf{x}}, \varepsilon)\mathcal{D}(\hat{\mathbf{y}}, \varepsilon) - \mathcal{D}(\hat{\mathbf{y}}, \varepsilon)\mathcal{D}(\hat{\mathbf{x}}, \varepsilon) = \mathcal{D}(\hat{\mathbf{z}}, \varepsilon^2) - 1 \quad (7.91)$$

up to $O(\varepsilon^4)$. This results in

$$-\frac{(J_x J_y - J_y J_x)\varepsilon^2}{\hbar^2} = \left(1 - \frac{iJ_z \varepsilon^2}{\hbar} \right) - 1 \quad (7.92)$$

which simplifies to

$$[J_x, J_y] = i\hbar J_z \quad (7.93)$$

By rotating and mirroring our coordinate system we can show

$$[J_i, J_j] = i\hbar \varepsilon_{ijk} J_k \quad (7.94)$$

7.3.1 Rotations of $\text{Spin-}\frac{1}{2}$ systems

For spin- $\frac{1}{2}$ systems the generators of rotations are

$$\begin{aligned} S_x &= \left(\frac{\hbar}{2} \right) [|+\rangle\langle -| + |- \rangle\langle +|], & S_y &= \left(\frac{i\hbar}{2} \right) [|- \rangle\langle +| - |+\rangle\langle -|], \\ S_z &= \left(\frac{\hbar}{2} \right) [|+\rangle\langle +| - |- \rangle\langle -|] \end{aligned} \quad (7.95)$$

Consider, for instance, rotations about the z-axis:

$$|\alpha'\rangle = \exp \left(- \left(\frac{iS_z}{\hbar} \right) \phi \right) |\alpha\rangle \quad (7.96)$$

Expanding $|\alpha\rangle = \langle +|\alpha\rangle |+\rangle + \langle -|\alpha\rangle |- \rangle$:

$$|\alpha'\rangle = \exp \left(- \frac{i\phi}{2} \right) |+\rangle \langle +|\alpha\rangle + \exp \left(\frac{i\phi}{2} \right) |- \rangle \langle -|\alpha\rangle \quad (7.97)$$

which implies that for rotations by $\phi = 2\pi$ we do not return to the original state, but rather

$$|\alpha\rangle \xrightarrow{\mathcal{D}(\hat{\mathbf{z}}, 2\pi)} -|\alpha\rangle \quad (7.98)$$

hence we require a 4π rotation to return back to the original state.

For instance the rotation of the expectation value of S_x :

$$\langle S'_x \rangle = \exp(-i\phi) \langle \alpha | - \rangle \langle - | S_x | + \rangle \langle + | \alpha \rangle \quad (7.99)$$

$$+ \exp(i\phi) \langle \alpha | + \rangle \langle + | S_x | - \rangle \langle - | \alpha \rangle \quad (7.100)$$

$$= \frac{\hbar}{2} (\exp(-i\phi) \langle \alpha | - \rangle \langle + | \alpha \rangle + \exp(i\phi) \langle \alpha | + \rangle \langle - | \alpha \rangle) \quad (7.101)$$

$$= \cos(\phi) \langle S_x \rangle - \sin(\phi) \langle S_y \rangle \quad (7.102)$$

Spin Precession

The Hamiltonian that describes a spin- $\frac{1}{2}$ particle in the presence of a magnetic field is

$$H = -\left(\frac{e}{mc}\right) \mathbf{S} \cdot \mathbf{B} = \omega S_z \quad (7.103)$$

hence the time evolution operator is

$$\mathcal{U}(t, 0) = \exp\left(-\frac{iS_z\omega t}{\hbar}\right) \quad (7.104)$$

which, as per our calculation above, tells us that

$$|\alpha(t)\rangle = \exp\left(-\frac{i\omega t}{2}\right) |+\rangle \langle +|\alpha_0\rangle + \exp\left(\frac{i\omega t}{2}\right) |-\rangle \langle -|\alpha_0\rangle \quad (7.105)$$

and that

$$\langle S_x \rangle_t = \langle S_x \rangle_{t=0} \cos(\omega t) - \langle S_y \rangle_{t=0} \sin(\omega t) \quad (7.106)$$

Pauli Formalism

The Pauli formalism expresses the state kets as spinors:

$$|+\rangle \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \chi_+, \quad |-\rangle \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \chi_- \quad (7.107)$$

hence a general spinor can be written as

$$\chi = \langle +|\alpha \rangle \chi_+ + \langle -|\alpha \rangle \chi_- = \begin{pmatrix} \langle +|\alpha \rangle \\ \langle -|\alpha \rangle \end{pmatrix} \quad (7.108)$$

In this basis the spin operators are given by the Pauli matrices:

$$S_x = \frac{\hbar}{2} \sigma_x, \quad S_y = \frac{\hbar}{2} \sigma_y, \quad S_z = \frac{\hbar}{2} \sigma_z \quad (7.109)$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (7.110)$$

Rotations in the Pauli Formalism

Using the Pauli matrices we can express the rotation operator in a simpler fashion

$$\exp\left(-\frac{i\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}\phi}{2}\right) = \mathbb{1} \cos\left(\frac{\phi}{2}\right) - i\boldsymbol{\sigma} \cdot \hat{\mathbf{n}} \sin\left(\frac{\phi}{2}\right) \quad (7.111)$$

which actually also works generally:

$$\exp\left(-\frac{i\mathbf{S} \cdot \hat{\mathbf{n}}\phi}{\hbar}\right) = \mathbb{1} \cos\left(\frac{\phi}{2}\right) - 2i\left(\frac{\mathbf{S} \cdot \hat{\mathbf{n}}}{\hbar}\right) \sin\left(\frac{\phi}{2}\right) \quad (7.112)$$

Using these rotations we can construct the eigenspinors of spin operators about an arbitrary axis $\hat{\mathbf{n}}$ (instead of the z -axis). This is achieved by rotating the system twice, once about the y axis (for instance), and once about the z -axis, with angles α and β respectively. Rotating χ_+ gives us

$$\begin{aligned} \chi'_+ &= \mathcal{D}(\hat{\mathbf{z}}, \beta) \mathcal{D}(\hat{\mathbf{y}}, \alpha) \chi_+ \\ &= \left(\mathbb{1} \cos\left(\frac{\beta}{2}\right) - i\sigma_z \sin\left(\frac{\beta}{2}\right) \right) \left(\mathbb{1} \cos\left(\frac{\alpha}{2}\right) - i\sigma_y \sin\left(\frac{\alpha}{2}\right) \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned} \quad (7.113)$$

resulting in

$$\chi'_+ = \begin{pmatrix} \cos\left(\frac{\beta}{2}\right) \exp\left(-\frac{i\alpha}{2}\right) \\ \sin\left(\frac{\beta}{2}\right) \exp\left(\frac{i\alpha}{2}\right) \end{pmatrix} \quad (7.114)$$

Euler Rotations

An arbitrary rotation can be accomplished by rotating three times:

$$R(\alpha, \beta, \gamma) = R_{z'}(\gamma) R_{y'}(\beta) R_z(\alpha) \quad (7.115)$$

where y' and z' refer to the rotated y and z axes. It can be shown that this arbitrary rotation can equivalently be accomplished by rotating about the original axes:

$$R(\alpha, \beta, \gamma) = R_z(\alpha) R_y(\beta) R_z(\gamma) \quad (7.116)$$

where the rotations now take place with respect to the fixed axis. Note that the angles have switched around! We can generalise this to our quantum mechanical rotation operators too:

$$\mathcal{D}(\alpha, \beta, \gamma) = \mathcal{D}_z(\alpha) \mathcal{D}_y(\beta) \mathcal{D}_z(\gamma) \quad (7.117)$$

which we can express as a matrix in the Pauli formalism:

$$\begin{aligned} &\exp\left(-\frac{i\sigma_z\alpha}{2}\right) \exp\left(-\frac{i\sigma_y\beta}{2}\right) \exp\left(-\frac{i\sigma_z\gamma}{2}\right) \\ &= \begin{pmatrix} e^{-\frac{i(\alpha+\gamma)}{2}} \cos\left(\frac{\beta}{2}\right) & e^{-\frac{i(\alpha-\gamma)}{2}} \sin\left(\frac{\beta}{2}\right) \\ e^{\frac{i(\alpha-\gamma)}{2}} \sin\left(\frac{\beta}{2}\right) & e^{\frac{i(\alpha+\gamma)}{2}} \cos\left(\frac{\beta}{2}\right) \end{pmatrix} \end{aligned} \quad (7.118)$$

7.3.2 Eigenvalues and Eigenstates of Angular Momentum

Ladder operators

The operators \mathbf{J} and J_z commute, therefore, there exists a simultaneous eigenbasis, which we will denote $|j, m\rangle$, for which

$$J^2|j, m\rangle = \hbar^2 j(j+1)|j, m\rangle, \quad J_z|j, m\rangle = \hbar m|j, m\rangle \quad (7.119)$$

Just like for the harmonic oscillator we can introduce ladder operators that take us from one state to one of the adjacent states:

$$J_{\pm} = J_x \pm iJ_y \quad (7.120)$$

For which

$$J_{\pm}|j\ m\rangle = \hbar\sqrt{(j \mp m)(j \pm m + 1)}|j, m \pm 1\rangle \quad (7.121)$$

Notice that this implies that any matrix representation of the rotation operator is block diagonal, where different values of j are not mixed with each other, which means it suffices to express the rotation of $|j\ m\rangle$ in terms of the different m values with total angular momentum unchanged.

$$\mathcal{D}(R)|j\ m\rangle = \sum_{m'} |j\ m'\rangle \mathcal{D}_{m' m}^{(j)}(R) \quad (7.122)$$

Generally

$$\mathcal{D}^{(j)}_{m' m}(\alpha, \beta, \gamma) = e^{-i(m' \alpha + m \gamma)} d_{m' m}^{(j)} \quad (7.123)$$

where

$$d_{m' m}^{(j)} = \langle j\ m' | \exp\left(-\frac{iJ_y \beta}{\hbar}\right) | j\ m \rangle \quad (7.124)$$

for $j = \frac{1}{2}$ this becomes

$$d^{\frac{1}{2}} = \begin{pmatrix} \cos\left(\frac{\beta}{2}\right) & -\sin\left(\frac{\beta}{2}\right) \\ \sin\left(\frac{\beta}{2}\right) & \cos\left(\frac{\beta}{2}\right) \end{pmatrix} \quad (7.125)$$

as we saw previously. Using this method it is quite easy to find the matrix representations of rotations for larger values of j .

Orbital Angular Momentum

When $j \in \mathbb{Z}$ the angular momentum operator, \mathbf{L} , can be written as

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} \quad (7.126)$$

Consider now the position representation of a state in spherical coordinates, $\langle \mathbf{x} | \alpha \rangle = \langle r, \theta, \phi | \alpha \rangle$. In this basis we can more easily handle rotations. Let us establish the effect of the angular momentum operators on these wavefunctions. For instance an infinitesimal rotation about the z-axis:

$$\langle r, \theta, \phi | \mathbb{1} - i\left(\frac{\delta\phi}{\hbar}\right) L_z | \alpha \rangle = \langle r, \theta, \phi - \delta\phi | \alpha \rangle \quad (7.127)$$

$$= \langle r, \theta, \phi | \alpha \rangle - \delta\phi \frac{\partial}{\partial\phi} \langle r, \theta, \phi | \alpha \rangle \quad (7.128)$$

which tells us that

$$L_z \langle r, \theta, \phi | \alpha \rangle = -i\hbar \frac{\partial}{\partial\phi} \langle r, \theta, \phi | \alpha \rangle \quad (7.129)$$

we can do similar calculations for L_x and L_y . Using these expression we can derive differential equations that describe $\langle r, \theta, \phi | \ell \ell \rangle$ (because using L_+ on this gives you zero). And from there we

can use the lowering operators on this wave function to get the remaining functions, as we did for the harmonic oscillator. The wavefunctions are given by the spherical harmonics:

$$\langle r, \theta, \phi | n \ell m \rangle = R_{n\ell}(r) Y_{\ell}^m(\theta, \phi) \quad (7.130)$$

though the radial part is not part of this discussion, so we should rather say

$$\langle \theta, \phi | \ell m \rangle = Y_{\ell}^m(\theta, \phi) \quad (7.131)$$

Schrödinger Equation for Central Potentials

Spherically symmetric potentials are symmetric under rotations. Thus a spherically symmetric potential will lead to a Hamiltonian which has spherically symmetric eigenfunctions: the eigenfunctions are given by the spherical harmonics. For spherical potentials we can thus find a basis in which H , \mathbf{L}^2 and L_z are diagonal simultaneous:

$$H|n \ell m\rangle = E|n \ell m\rangle \quad (7.132)$$

$$\mathbf{L}^2|n \ell m\rangle = \hbar^2 \ell(\ell + 1)|n \ell m\rangle \quad (7.133)$$

$$L_z|n \ell m\rangle = \hbar m|n \ell m\rangle \quad (7.134)$$

We already know that the angular part of the wavefunctions is given by the spherical harmonics. We now are looking for the radial part, which is described by the radial equation:

$$\left(-\frac{\hbar^2}{2mr^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) + \frac{\hbar^2 \ell(\ell + 1)}{2mr^2} + V(r) \right) R_{n\ell}(r) = ER_{n\ell} \quad (7.135)$$

Using this we can solve the Schrödinger Equation for the Hydrogen atom, and hydrogenic atoms alike. It also serves as an unperturbed Hamiltonian in the Helium atom for instance.

Addition of Angular Momentum

Now we will concern ourselves with coupled angular momenta. The total angular momentum can be written as

$$\mathbf{J} = \mathbf{J}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \mathbf{J}_2 \quad (7.136)$$

due to some form of coupling the uncoupled states $|j_1 m_1 j_2 m_2\rangle$ are no longer sufficient and we wish to find an eigenbasis of $\{J^2, J_z, J_1^2, J_2^2\}$. This is done using Clebsch-Gordon decomposition. Consider, for instance the maximal state $|j_1 j_1 j_2 j_2\rangle$, this *must* be equal to the maximal angular momentum state in the new basis too: $|j_1 j_2 j j\rangle$ from here we can use the lowering operator to find other states with the same value of j , and thereafter use orthonormality to find the other Hilbert-subspaces.

7.4 SYMMETRY IN QUANTUM MECHANICS

Let us define a general symmetry operator, \hat{S} , which is unitary because we need probability conservation. If the symmetry parametre is infinitesimal the first order expansion of the symmetry operator suffices:

$$\hat{S} = \mathbb{1} - \frac{i\varepsilon}{\hbar} G \quad (7.137)$$

where G is the hermitian generator of the symmetry. In cases where $[G, H] = 0$ we also have $[\mathcal{J}, H] = 0$ in which case G is a constant of motion. This allows us to find a basis of states that are eigenstates of both H and G .

7.4.1 Symmetry

Consider, for example, a rotationally invariant Hamiltonian. We know that

$$[\mathbf{J}, H] = \mathbf{0}, \quad [J^2, H] = 0 \quad (7.138)$$

hence there exists a basis that is diagonal in $\{H, J^2, J_z\}$, which we once again denote $|n j m\rangle$. Due to the rotational invariance we know that

$$H\mathcal{D}(R)|n j m\rangle = \mathcal{D}(R)H|n j m\rangle \quad (7.139)$$

hence we know that the unrotated and the rotated kets have the same energy. As we saw previously rotations mix different values of m together (all with the same value of j), therefore we can conclude that $\{|n, j, -j\rangle, |n, j, (1-j)\rangle, \dots, |n, j, (j-1)\rangle, |n, j, j\rangle\}$ forms a degenerate space. Alternatively we see this through the fact that $[J_{\pm}, H] = 0$.

This tells us that the degeneracy can only be broken by a perturbation that is *not* rotationally invariant, which we saw for the Hydrogen atom when we apply a magnetic field.

7.4.2 Discrete Symmetries

Until now we have only concerned ourselves with continuous symmetry operations; ones for which we could make an infinitesimal expansion. However, there exist symmetry operations that are discrete the first of which is the *parity* operation.

Parity

The parity operation, also called space-inversion, is represented by a unitary operator, π , which inverts all of our spatial coordinates:

$$\langle \alpha | \pi^\dagger \mathbf{x} \pi | \alpha \rangle = -\langle \alpha | \mathbf{x} | \alpha \rangle \quad (7.140)$$

which tells us that the parity operator also inverts our spatial eigenkets:

$$\pi | \mathbf{x} \rangle = e^{i\delta} | -\mathbf{x} \rangle \quad (7.141)$$

up to some phase factor. Let us look at infinitesimal translations:

$$\pi \mathcal{T}(\mathbf{dx}) = \mathcal{T}(-\mathbf{dx}) \pi \quad (7.142)$$

so

$$\pi \left(\mathbb{1} - \frac{i\mathbf{p} \cdot \mathbf{dx}}{\hbar} \right) \pi^\dagger = \left(\mathbb{1} + \frac{i\mathbf{p} \cdot \mathbf{dx}}{\hbar} \right) \quad (7.143)$$

and so

$$\pi^\dagger \mathbf{p} \pi = -\mathbf{p} \quad (7.144)$$

this makes intuitive sense: if we invert our coordinate systems things moving in direction $\hat{\mathbf{n}}$ will move in direction $-\hat{\mathbf{n}}$ in our new coordinate system.

However, rotations commute with the parity operator, hence

$$\pi^\dagger \mathbf{J} \pi = \mathbf{J} \quad (7.145)$$

This implies

$$\pi^\dagger \mathbf{S} \cdot \mathbf{x} \pi = -\mathbf{S} \cdot \mathbf{x} \quad (7.146)$$

and

$$\pi^\dagger \mathbf{L} \cdot \mathbf{S} \pi = \mathbf{L} \cdot \mathbf{S} \quad (7.147)$$

Parity of wavefunctions

Suppose $[H, \pi] = 0$, then we can find mutual eigenstates, therefore, it must be the case that the eigenstates are either even or odd under parity operations.

For instance the harmonic oscillator has a Hamiltonian that is invariant under parity, which is why we know that its eigenfunctions are either symmetric or anti-symmetric under spatial inversion.

Parity selection rule

Consider eigenkets of a Hamiltonian that commutes with π . These eigenkets must be parity eigenkets,

$$\pi|i\rangle = \varepsilon_i|i\rangle, \quad \varepsilon_i^2 = 1 \quad (7.148)$$

This tells us the matrix elements of (for instance) x are zero unless the initial and final states have opposite parity.

$$\langle f|\mathbf{x}|i\rangle = \langle f|\pi^\dagger \pi \mathbf{x} \pi^\dagger \pi|i\rangle = -\varepsilon_i \varepsilon_f \langle f|\mathbf{x}|i\rangle \quad (7.149)$$

Lattice Translations

Consider a periodic potential, with lattice constant a :

$$\mathcal{T}^\dagger(a)V(x)\mathcal{T}(a) = V(x+a) = V(x) \quad (7.150)$$

therefore we should expect that the energy eigenfunctions also are invariant under (discrete) translations. We can construct the ground state by looking at the ground state of the n^{th} unit cell. This is obviously not an energy eigenstate, as we have $\mathcal{T}(a)|n\rangle = e^{i\delta}|n+1\rangle$. Thus the translation

operator relates every ground state to its neighbour's ground state: the total ground state must be a linear combination of all these:

$$|\theta\rangle = \sum_n e^{in\theta}|n\rangle \quad (7.151)$$

for this we have

$$\mathcal{T}(a)|\theta\rangle = e^{i\delta} \sum_n e^{in\theta}|n+1\rangle = e^{i(\delta-\theta)}|\theta\rangle \quad (7.152)$$

we choose $\delta = 0$ by convention.

To evaluate the energy of this state let us assume that

$$\langle n'|H|n\rangle = E_0\delta_{n,n'} - \Delta\delta_{n,n'+1} - \Delta\delta_{n,n'-1} \quad (7.153)$$

this is the tight-binding approximation. Then

$$E_\theta = E_0 - 2\Delta \cos \theta \quad (7.154)$$

so the ground state is $|\theta = 0\rangle$. Bloch's theorem states that a periodic potential with lattice constant a has eigenstates of the form

$$\psi_k(x) = u_k(x)e^{ikx} \quad (7.155)$$

where $u_k(x) = u_k(x+a)$, which tells us that the angle kx is related to the wavenumber, k : the solution to this problem is hence a slight (periodic) modification to planar waves.

Time-Reversal Symmetry

The next discrete symmetry operation is time reversal, which is often also thought of as motion inversion. Kepler's problem is invariant under time-reversal, if a trajectory $\mathbf{r}(t)$ is a solution to Kepler's equations, then so is $\mathbf{r}(-t)$. The operator, Θ , that represents time-reversal in quantum mechanics is *antiunitary*:

Antiunitary operators still preserve the norm, but they do not preserve inner products, due to the complex conjugation

$$\Theta = U_\Theta K \quad (7.156)$$

where U_Θ is unitary and K is the complex conjugation operator. The antiunitarity is a necessity, otherwise we would get that time-reversed energy eigenstate have negative energies. This is a new kind of operator, which has some weird properties, for instance let

Isn't that exactly what we said for QFT?

$$|\tilde{\alpha}\rangle = \Theta|\alpha\rangle, \quad |\tilde{\beta}\rangle = \Theta|\beta\rangle \quad (7.157)$$

Then

$$\langle \tilde{\alpha} | \tilde{\beta} \rangle = \langle \beta | \alpha \rangle \quad (7.158)$$

NB: The time-reversal operator only preserves the magnitude of inner products.

For a linear operator, X :

$$\langle \alpha | X | \beta \rangle = \langle \tilde{\alpha} | \Theta X^\dagger \Theta^{-1} | \tilde{\beta} \rangle \quad (7.159)$$

We say that an observable is odd or even under time reversal according to we have the upper or lower sign in

$$\Theta A \Theta^{-1} = \pm A \quad (7.160)$$

Momentum is odd under time-reversal, as we require, and position is even

$$\Theta \mathbf{p} \Theta^{-1} = -\mathbf{p}, \quad \Theta \mathbf{x} \Theta^{-1} = \mathbf{x} \quad (7.161)$$

Angular momentum is odd

$$\Theta \mathbf{J} \Theta^{-1} = -\mathbf{J} \quad (7.162)$$

Wavefunctions

Consider a state in its position representation

$$|\alpha\rangle = \int d^3x |\mathbf{x}\rangle \langle \mathbf{x}|\alpha\rangle \quad (7.163)$$

Applying time-reversal

$$\Theta |\alpha\rangle = \int d^3x |\mathbf{x}\rangle \langle \alpha|\mathbf{x}\rangle \quad (7.164)$$

which tells us that

$$\psi(\mathbf{x}) \xrightarrow{\Theta} \psi^*(\mathbf{x}) \quad (7.165)$$

This can also be seen by looking at the Schrödinger Equation

$$i\hbar \partial_t \psi(\mathbf{x}) = \hat{H} \psi(\mathbf{x}) \quad (7.166)$$

(for H invariant under time-reversal). Applying Θ

$$-i\hbar \partial_t (\Theta \psi(\mathbf{x})) = \hat{H} (\Theta \psi(\mathbf{x})) \quad (7.167)$$

But by simply taking the complex conjugate of the Schrödinger Equation we obtain the same differential equation

$$-i\hbar \partial_t \psi^*(\mathbf{x}) = \hat{H} \psi^*(\mathbf{x}) \quad (7.168)$$

thus

$$\psi(-t, \mathbf{x}) = \psi^*(t, \mathbf{x}) \quad (7.169)$$

Time-reversal of spin- $\frac{1}{2}$ systems

Consider the eigenket of $\mathbf{S} \cdot \hat{\mathbf{n}}$ with eigenvalue $\frac{\hbar}{2}$

$$|\hat{\mathbf{n}}, +\rangle = \exp\left(-\frac{iS_z\alpha}{\hbar}\right) \exp\left(-\frac{iS_y\beta}{\hbar}\right) |+\rangle \quad (7.170)$$

and we know that

$$\Theta|\hat{\mathbf{n}}, +\rangle = \eta|\hat{\mathbf{n}}, -\rangle \quad (7.171)$$

for some phase factor η . We can also construct this state through

$$|\hat{\mathbf{n}}, -\rangle = \exp\left(-\frac{iS_z\alpha}{\hbar}\right) \exp\left(-\frac{iS_y(\beta + \pi)}{\hbar}\right) |+\rangle \quad (7.172)$$

hence

$$\Theta = \eta \exp\left(-\frac{i\pi S_y}{\hbar}\right) K \quad (7.173)$$

and in fact this holds generally

$$\Theta = \eta \exp\left(-\frac{i\pi J_y}{\hbar}\right) K \quad (7.174)$$

which we can use to show that

$$\Theta^2 = (-1)^{2j} \quad (7.175)$$

Kramer's Degeneracy

For half-integer-spin systems where $[H, \Theta] = 0$ we have that

$$\Theta^2|n\rangle = -|n\rangle \quad (7.176)$$

assume now that E_n is non-degenerate, then

$$\Theta|n\rangle = e^{i\delta}|n\rangle, \quad \Theta^2|n\rangle = |n\rangle \quad (7.177)$$

however, this contradicts our assumption. Therefore we *know* that Hamiltonians that are invariant under time-reversal have at least two-fold degeneracies at each energy level, assuming they describe a half-integer spin system

7.5 APPROXIMATION METHODS

We will now consider Hamiltonians that cannot be solved analytically and must be approximated.

7.5.1 *Perturbation Theory: Non-Degenerate*

The first approximation method we will look at is perturbation theory, where we split Hamiltonian up, into an unperturbed Hamiltonian, H_0 , which we can solve analytically and a perturbation V . However, it is standard procedure to introduce a perturbation parametre λ which is used to keep track of the order of the perturbation:

$$H = H_0 + \lambda V \quad (7.178)$$

As mentioned, we know the eigenkets and eigenenergies of H_0 :

$$H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \quad (7.179)$$

The effect of the perturbation is now written as a power series:

$$|n\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \dots \quad (7.180)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (7.181)$$

where we refer to the n^{th} term in the sum as the n^{th} order correction. Up to second order we have

$$E_n = E_n^{(0)} + \lambda \langle n^{(0)}|V|n^{(0)}\rangle + \lambda^2 \sum_{k \neq n} \frac{|\langle n^{(0)}|V|k^{(0)}\rangle|^2}{E_n^{(0)} - E_k^{(0)}} \quad (7.182)$$

and

$$|n\rangle = |n^{(0)}\rangle + \lambda \sum_{k \neq n} \frac{\langle n^{(0)}|V|k^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}} |k^{(0)}\rangle \quad (7.183)$$

$$+ \lambda^2 \left(\sum_{k, \ell \neq n} \frac{|k^{(0)}\rangle V_{k\ell} V_{\ell n}}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_\ell^{(0)})} - \sum_{k \neq n} \frac{|k^{(0)}\rangle V_{nn} V_{k\ell}}{(E_n^{(0)} - E_k^{(0)})^2} \right) \quad (7.184)$$

where

$$V_{ij} = \langle i^{(0)}|V|j^{(0)}\rangle \quad (7.185)$$

Note that $\langle n^{(0)}|n^{(0)}\rangle = 1$, therefore these wavefunctions are not normalized. However, using the formulae above it is quite easy to show that the normalisation constant $\sqrt{Z_n}$ can be found through

$$Z_n^{-1} = \langle n|n\rangle = 1 + \lambda^2 \sum_{k \neq n} \frac{|V_{kn}|^2}{(E_n^{(0)} - E_k^{(0)})^2} + O(\lambda^3) \quad (7.186)$$

Harmonic Oscillator

Let us look at the harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \quad (7.187)$$

And we will add a perturbation $\varepsilon V = \frac{1}{2}\varepsilon m\omega^2 x^2$. We can quite easily find the matrix elements of the perturbation using the raising and lowering operators:

$$V_{00} = \varepsilon \frac{\hbar\omega}{4} \quad (7.188)$$

$$V_{20} = \varepsilon \frac{\hbar\omega}{2\sqrt{2}} \quad (7.189)$$

These are the only two terms that contribute to the first order correction of the ground state

$$|0\rangle = |0^{(0)}\rangle - \frac{\varepsilon}{4\sqrt{2}}|2^{(0)}\rangle + O(\varepsilon^2) \quad (7.190)$$

and

$$E_0 = \frac{\hbar\omega}{2} \left(1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + O(\varepsilon^3) \right) \quad (7.191)$$

which coincides to the first three terms in the full solution, which we can calculate exactly $E_0 = \left(\frac{\hbar\omega}{2}\right) \sqrt{1 + \varepsilon}$. Naturally this approach also holds in position space, therefore the terms in Equation tell us about the corrections to the wavefunction. Notices that the potential is still even under parity, therefore the perturbed wavefunctions will still either be odd or even under parity, which is why $|0^{(0)}\rangle$ is mixed with $|2^{(0)}\rangle$ and *not* $|1^{(0)}\rangle$

7.5.2 Perturbation Theory: Degenerate

In the case the we have degenerate eigenvalues we cannot use the previous expressions, because even though there are two kets $|n\rangle$ and $|k\rangle$, we cannot know that they do not have the same energy. However, the perturbation often removes the degeneracy, which means that every ket in the degenerate space has their own energy. Thus, there exists a basis within the degenerate space such that the perturbation is diagonal. Therefore in degenerate perturbation theory, we look exclusively inside degenerate space and perform the perturbation theory in the appropriate basis, which is found by finding the eigenkets of V .

Interestingly, the second order correction from within the degenerate space is zero, so the second order correction is

$$E_\ell^{(2)} = \sum_{k \notin D} \frac{|V_{k\ell}|^2}{E_D^{(0)} - E_k^{(0)}} \quad (7.192)$$

where D is the *degenerate subspace*; the sum only goes over kets that are not degenerate with $|\ell^{(0)}\rangle$, hence we need not worry about the fraction exploding.

7.5.3 Perturbation Theory: Time-Dependent

Consider now a time dependent perturbation

$$H = H_0 + V(t)\Theta(t - t_0) \quad (7.193)$$

Where we once again know the solution to H_0 . We wish to find $c_n(t)$ for $t > t_0$ such that

$$|\alpha, t_0; t\rangle = \sum_n c_n(t) e^{-i\frac{E_n t}{\hbar}} |n\rangle \quad (7.194)$$

Interaction Picture

The interaction picture is useful to time-dependent perturbation theory, as we will see. The interaction picture state ket is defined by

$$|\alpha, t_0; t\rangle_I = e^{iH_0 t/\hbar} |\alpha, t_0; t\rangle_S \quad (7.195)$$

Operators are

$$A_I \equiv e^{iH_0 t/\hbar} A_S e^{-iH_0 t/\hbar} \quad (7.196)$$

Using these definitions we can show that the Schrödinger Equation becomes

$$i\hbar \partial_t |\alpha, t_0; t\rangle_I = V_I |\alpha, t_0; t\rangle_I \quad (7.197)$$

The time-evolution is only due to the (time-dependent) perturbation! On the other hand, the time-evolution of operators is due to the unperturbed Hamiltonian only

$$\frac{dA_I}{dt} = \frac{1}{i\hbar} [A_I, H_0] \quad (7.198)$$

We can rewrite the Schrödinger Equation for kets such that it includes the terms $c_n(t)$ mentioned above, in order to do so we expand using the eigenkets of the unperturbed Hamiltonian

$$i\hbar \partial_t \langle n|\alpha, t_0; t\rangle_I = \sum_m \langle n|V_I|m\rangle \langle m|\alpha, t_0; t\rangle_I \quad (7.199)$$

but

$$\langle n|V_I|m\rangle = \langle n|e^{iE_n t/\hbar} V_S e^{-iE_m t/\hbar}|m\rangle = e^{-i(E_m - E_n)t/\hbar} V_{nm}(t) \quad (7.200)$$

and naturally $\langle n|\alpha, t_0; t\rangle = c_n(t)$. Thus we can write this as a matrix differential equation with time-dependent coefficients

$$i\hbar \partial_t \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12}e^{i\omega_{12}t} & \cdots \\ V_{21}e^{i\omega_{21}t} & V_{22} & \cdots \\ & & V_{33} & \cdots \\ \vdots & \cdots & & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} \quad (7.201)$$

For instance, consider a two-level system with Hamiltonian

$$H = \begin{pmatrix} E_1 & \gamma e^{i\omega t} \\ \gamma e^{-i\omega t} & E_2 \end{pmatrix} \quad (7.202)$$

clearly

$$i\hbar \begin{pmatrix} \dot{c}_1 \\ \dot{c}_2 \end{pmatrix} = \begin{pmatrix} 0 & \gamma e^{i(\omega - \omega_0)t/\hbar} \\ \gamma e^{-i(\omega - \omega_0)t/\hbar} & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (7.203)$$

We can decouple these to a second order differential equation with constant coefficients

$$\ddot{c}_1 = i\delta \dot{c}_1 - \left(\frac{\gamma}{\hbar}\right)^2 c_1, \quad \delta = \omega - \omega_0 \quad (7.204)$$

which can be solved using a Laplace Transformation. The result is given by *Rabi's Formula*

$$|c_2(t)|^2 = \frac{\gamma^2/\hbar^2}{\gamma^2/\hbar^2 + \delta^2/4} \sin^2 \left[\left(\frac{\gamma^2}{\hbar^2} + \frac{\delta^2}{4} \right)^{\frac{1}{2}} t \right] \quad (7.205)$$

$$|c_1(t)|^2 = 1 - |c_2(t)|^2 \quad (7.206)$$

(the initial condition used to solve this is $c_1(0) = 1$)

Dyson Series

Seeing as the Schrödinger Equation only depends on the perturbation, we can write the time-evolution operator only as a function of $V_I(t)$ too:

$$U_I(t, t_0) = \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt' V_I(t') + \left(\frac{-i}{\hbar} \right)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' V_I(t') V_I(t'') + \dots \quad (7.207)$$

Transition Probability

Using the Dyson Series one can derive a perturbative expression for the transition probability:

$$c_f(t) = \langle f | U_I(t, t_0) | i \rangle \quad (7.208)$$

so the first and second order terms would be

$$c_f^{(1)} = -\frac{i}{\hbar} \int_{t_0}^t dt' \langle f | V_I(t') | i \rangle = -\frac{i}{\hbar} \int_{t_0}^t dt' e^{i\omega_{fi}t'} V_{fi}(t') \quad (7.209)$$

$$c_f^{(2)} = \left(\frac{-i}{\hbar} \right)^2 \sum_m \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' e^{i(\omega_{fm}t' + \omega_{mi}t'')} V_{fm}(t') V_{mi}(t'') \quad (7.210)$$

Let us, for instance, consider a constant perturbation

$$H = H_0 + V \Theta(t - t_0) \quad (7.211)$$

here, for $t_0 = 0$

$$c_f^{(1)} = -\frac{i}{\hbar} V_{fi} \int_0^t dt' e^{i\omega_{fi}t'} \quad (7.212)$$

$$= \frac{V_{fi}}{E_f - E_i} (1 - e^{i\omega_{fi}t}) \quad (7.213)$$

which means

$$|c_f^{(1)}|^2 = \frac{|V_{fi}|^2}{|E_f - E_i|^2} \sin^2 \left[\frac{(E_f - E_i)t}{2\hbar} \right] \quad (7.214)$$

7.5.4 Fermi's Golden Rule

In cases where we have a continuous energy spectrum and are looking at transitions where $E_f \approx E_i$ we can replace a sum such as $\sum_{f, E_f \approx E_i} |c_f^{(1)}|^2$ with

$$\sum_{f, E_f \approx E_i} |c_f^{(1)}|^2 \rightsquigarrow \int dE_f \rho(E_f) |c_f^{(1)}|^2 \quad (7.215)$$

where the density of states $\rho(E)dE$ is defined as the number of states within the energy interval $(E, E + dE)$. Using this we can define the transition rate

$$w_{i \rightarrow [f]} = \frac{2\pi}{\hbar} \overline{|V_{fi}|^2} \rho(E_f)_{E_f \approx E_i} \quad (7.216)$$

where $[f]$ is the set of final states in the neighbourhood of $|i\rangle$. We sometimes also write this expression as

$$w_{i \rightarrow f} = \frac{2\pi}{\hbar} |V_{fi}|^2 \delta(E_n - E_i) \quad (7.217)$$

Where it must be understood that this expression is integrated with the measure $\int dE_f \rho(E_f)$

7.5.5 Harmonic Perturbation

Consider now a harmonic perturbation of the form

$$V(t) = (\mathcal{U}e^{i\omega t} + \mathcal{U}^\dagger e^{-i\omega t}) \Theta(t - 0) \quad (7.218)$$

for \mathcal{U} some operator that may further depend on, position, momentum, spin etc. Using our time-dependent perturbation theory we can write

$$c_n^{(1)} = -\frac{i}{\hbar} \int_0^t dt' (\mathcal{U}_{ni} e^{i\omega t'} + \mathcal{U}_{ni}^\dagger e^{-i\omega t'}) e^{i\omega_{ni} t'} \quad (7.219)$$

$$= \frac{1}{\hbar} \left[\frac{1 - e^{i(\omega + \omega_{ni})t}}{\omega + \omega_{ni}} - \frac{1 - e^{i(\omega - \omega_{ni})t}}{\omega - \omega_{ni}} \right] \quad (7.220)$$

This looks a lot like the term we saw for the constant perturbation, if we just let $\omega_{ni} \rightsquigarrow \omega_{ni} \pm \omega$. The transition probability is only appreciable if $\omega_{ni} \pm \omega \approx 0$. These two situations correspond to stimulated emission and absorption respectively. Note that we have assumed that the perturbation is classical, so we have not limited the amount of energy it can give away or absorb; we can freely take as much energy as needed from a bath, but also gives as much energy as needed to the bath. Here the transition amplitudes become

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} \begin{cases} |V_{ni}|^2 \delta(E_n - E_i + \hbar\omega) \\ |V_{ni}^\dagger|^2 \delta(E_n - E_i - \hbar\omega) \end{cases} \quad (7.221)$$

however, naturally $|V_{ni}| = |V_{ni}^\dagger|$ which tells us that if the density of states is constant we have as much emission as we do absorption. However, if the density of states is different we get

$$\frac{w_{E_i \rightarrow E_i - \hbar\omega}}{\rho(E_i - \hbar\omega)} = \frac{w_{E_i \rightarrow E_i + \hbar\omega}}{\rho(E_i + \hbar\omega)} = \frac{2\pi}{\hbar} \quad (7.222)$$

7.5.6 Energy Shift and Decay Width

Consider a constant perturbation, V , which is turned on very slowly $e^{\eta t} V$. At $t = -\infty$ the perturbation is switched off and it gradually gets turned on. In the limit $\eta \rightarrow 0$ we retrieve our original constant perturbation.

We are now interested in the coefficient $c_i(t)$, that is, we want to know how the population of $|i\rangle$ disappears as its energy is dispersed throughout the system.

$$c_i^{(1)} = -\frac{i}{\hbar} V_{ii} \int_{-\infty}^t dt' e^{\eta t'} = -\frac{i}{\hbar \eta} V_{ii} e^{\eta t} \quad (7.223)$$

$$c_i^{(2)} = \left(\frac{-i}{\hbar}\right)^2 |V_{ii}|^2 \frac{e^{2\eta t}}{2\eta^2} + \left(\frac{-i}{\hbar}\right) \sum_{m \neq i} \frac{|V_{mi}|^2 e^{2\eta t}}{2\eta(E_i - E_m + i\hbar\eta)} \quad (7.224)$$

Let us choose the Ansatz $c_i(t) = e^{-i\Delta_i t/\hbar}$; we expect the population of the initial state to decay exponentially and to have an oscillatory part. Using our first two corrections to c_i and some fancy maths, we get that

$$\Delta_i^{(2)} = \left(\text{Pr} \sum_{m \neq i} \frac{|V_{mi}|^2}{E_i - E_m} \right) - i \left(\pi \sum_{m \neq i} |V_{mi}|^2 \delta(E_i - E_m) \right) \quad (7.225)$$

The population indeed dies exponentially:

$$|c_i^{(2)}|^2 = e^{-\Gamma_i t/\hbar}, \quad \Gamma_i = 2\text{Im}(\Delta_i^{(2)}) \quad (7.226)$$

We can interpret this Γ_i as a decay *width*.

7.6 SCATTERING THEORY

Consider a Hamiltonian of the form

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}) \quad (7.227)$$

the unperturbed energy-levels are

$$E_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m} \quad (7.228)$$

The eigenvectors are plane waves, which we denote by $|\mathbf{k}\rangle$. Once again the transition amplitude is $\langle n|U_I(t, t_0)|i\rangle$ which is

$$\begin{aligned} \langle n|U_I(t, t_0)|i\rangle &= \mathbb{1} \delta_{ni} \\ &\quad - \frac{i}{\hbar} \sum_m V_{nm} \int_{t_0}^t dt' e^{i\omega_{nm}t'} \langle m|U_I(t', t_0)|i\rangle \end{aligned} \quad (7.229)$$

Let us define a matrix T , such that

$$\langle n|U_I(t, t_0)|i\rangle = \mathbb{1} \delta_{ni} - \frac{i}{\hbar} T_{ni} \int_{t_0}^t dt' e^{(i\omega_{ni} + \varepsilon)t'} \quad (7.230)$$

with $0 < \varepsilon \ll t^{-1}$. Now also define the scattering matrix, S , such that

$$S_{ni} \equiv \lim_{t \rightarrow \infty} \left(\lim_{\varepsilon \rightarrow 0} \langle n|U_I(t, -\infty)|i\rangle \right) \quad (7.231)$$

$$= \delta_{ni} - 2\pi i \delta(E_n - E_i) T_{ni} \quad (7.232)$$

Note that it is important that we take the ε limit first!

Because the perturbation is time-independent we can evaluate the integral

$$c_n(t) = \langle n | U_I(t, -\infty) | i \rangle = -\frac{i}{\hbar} T_{ni} \left(\frac{e^{(i\omega_{ni} + \varepsilon)t}}{i\omega_{ni} + \varepsilon} \right) \quad (7.233)$$

and hence the transition rate

$$w_{i \rightarrow n} = \frac{d}{dt} |c_n(t)|^2 = \frac{1}{\hbar^2} |T_{ni}|^2 \frac{2\varepsilon e^{2\varepsilon t}}{\omega_{ni}^2 + \varepsilon^2} \quad (7.234)$$

Taking the limit as $\varepsilon \rightarrow 0$ gives us $\sim \delta(\omega_{ni})$, because the fraction involving ε goes to zero in this limit unless ω_{ni} is zero, which again enforces energy conservation:

$$\lim_{\varepsilon \rightarrow 0} \frac{2\varepsilon e^{2\varepsilon t}}{\omega_{ni}^2 + \varepsilon^2} = 2\pi\hbar \delta(E_n - E_i) \quad (7.235)$$

giving us Fermi's Golden Rule once again

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} |T_{ni}|^2 \delta E_n - E_i \quad (7.236)$$

Consider now free particles confined inside a box of dimension L^3 . We can always let $L \rightarrow \infty$ again to return to the fully free particles. Now due to our boundary conditions we get that

$$E_{\mathbf{n}} = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L} \right)^2 (\mathbf{n})^2 \quad (7.237)$$

and hence

$$\rho(E_n) = \frac{\Delta n}{\Delta E_n} = \frac{mk}{\hbar} \left(\frac{L}{2\pi} \right)^2 d\Omega \quad (7.238)$$

where Ω is the solid angle in \mathbf{n} -space. The flux is $\mathbf{j} = \frac{\hbar \mathbf{k}}{mL^3}$ and the transition rate is

$$w_{i \rightarrow n} = \frac{mkL^3}{(2\pi)^2 \hbar^3} |T_{ni}|^2 d\Omega \quad (7.239)$$

The cross section times transition rate must be equal to the flux, hence we can isolate for the cross section (per steradian)

$$\frac{d\sigma}{d\Omega} = \left(\frac{mL^3}{2\pi\hbar^2} \right)^2 |T_{ni}|^2 \quad (7.240)$$

7.6.1 Finding the T-Matrix

It is by no means straightforward to find T , because in the only equation we have to describe it we cannot directly isolate it:

$$T_{ni} = V_{ni} + \sum_m V_{nm} \frac{T_{mi}}{E_i - E_m + i\hbar\varepsilon} \quad (7.241)$$

However we can expand this equation and write it in such a way that it is clear which terms is of order $O(V^n)$:

$$T = V \left(1 + \sum_n \left(\frac{1}{E_i - H_0 + i\hbar\varepsilon} V \right)^n \right) \quad (7.242)$$

note that the *operator* H_0 appears in the denominator therefore we need to be careful when we multiply the terms out, for instance the third order term is

$$V \frac{1}{E_i - H_0 + i\hbar\epsilon} V \frac{1}{E_i - H_0 + i\hbar\epsilon} V \neq V^3 \left(\frac{1}{E_i - H_0 + i\hbar\epsilon} \right)^2 \quad (7.243)$$

this is because in general $[H_0, V] \neq 0$. Equation 7.242 was derived through the *Lippmann-Schwinger Equation*, which defines a state, $|\psi^\pm\rangle$:

$$|\psi^\pm\rangle = |i\rangle + \frac{1}{E_i - H_0 \pm i\hbar\epsilon} V |\psi^\pm\rangle \quad (7.244)$$

which we will use in the following

7.6.2 Scattering Amplitude

Consider the position-space representation of $|\psi^\pm\rangle$:

$$\langle \mathbf{x} | \psi^\pm \rangle = \langle \mathbf{x} | i \rangle + \int d\mathbf{x}' \langle \mathbf{x} | \frac{1}{E_i - H_0 \pm i\hbar\epsilon} | \mathbf{x}' \rangle \langle \mathbf{x}' | V | \psi^\pm \rangle \quad (7.245)$$

$[H_0, \mathbf{x}] \neq 0$ but $[H_0, \mathbf{k}]$, so it is easiest to just to evaluate the Green's function, G_\pm in the momentum basis

$$G_\pm(\mathbf{x}, \mathbf{x}') = \frac{1}{(2\pi)^3} \int d\mathbf{k}' \frac{e^{i\mathbf{k}' \cdot (\mathbf{x} - \mathbf{x}')}}{k^2 - k'^2 \pm i\epsilon} \quad (7.246)$$

we can evaluate this with the Residue theorem:

$$G_\pm(\mathbf{x}, \mathbf{x}') = -\frac{1}{4\pi} \frac{e^{\pm i k |\mathbf{x} - \mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} \quad (7.247)$$

So

$$\langle \mathbf{x} | \psi^\pm \rangle = \langle \mathbf{x} | i \rangle - \frac{2m}{\hbar^2} \int d\mathbf{x}' \frac{e^{\pm i k |\mathbf{x} - \mathbf{x}'|}}{4\pi |\mathbf{x} - \mathbf{x}'|} V(\mathbf{x}') \langle \mathbf{x}' | \psi^\pm \rangle \quad (7.248)$$

The value $k |\mathbf{x} - \mathbf{x}'|$ is rather odd, so let us approximate it, in the limit where $r \gg r'$ where $r = |\mathbf{x}|$ and $r' = |\mathbf{x}'|$.

$$|\mathbf{x} - \mathbf{x}'| = \sqrt{r^2 - 2rr' \cos \alpha + r'^2} \approx r - \hat{\mathbf{r}} \cdot \mathbf{x}' \quad (7.249)$$

We get

$$\langle \mathbf{x} | \psi^+ \rangle \xrightarrow{\text{large } r} \frac{1}{L^{\frac{3}{2}}} \left[e^{i\mathbf{k} \cdot \mathbf{x}} + \frac{e^{ikr}}{r} f(\mathbf{k}, \mathbf{k}') \right] \quad (7.250)$$

thus, at large r (large distance from the localised potential) we can approximate the outgoing wave as the ingoing wave plus a *spherical* wave with amplitude $f(\mathbf{k}, \mathbf{k}')$. We have assumed that $|i\rangle = |\mathbf{k}\rangle$, i.e. that we sent planar waves into the sample.

$$f(\mathbf{k}, \mathbf{k}') = -\frac{mL^3}{2\pi\hbar^2} \langle \mathbf{k}' | V | \psi^+ \rangle \quad (7.251)$$

We refer to $f(\mathbf{k}, \mathbf{k}')$ as the *scattering amplitude*. The differential cross section becomes

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k}, \mathbf{k}')|^2 \quad (7.252)$$

7.6.3 Born Approximation

Let us consider again Equation 7.242:

$$T = V \left(1 + \sum_n \left(\frac{1}{E_i - H_0 + i\hbar\epsilon} V \right)^n \right) \quad (7.253)$$

Taking this up to first order, where $T = V$ is the (first order) *Born Approximation*:

$$f^{(1)}(\mathbf{k}, \mathbf{k}') = -\frac{m}{2\pi\hbar^2} \int d\mathbf{x} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} V(\mathbf{x}) \quad (7.254)$$

for the finite spherical well of radius a and height V_0 we get

$$f^{(1)}(\theta) = -\frac{2m}{\hbar^2} \frac{V_0 a}{(qa)^3} \left[\frac{\sin qa}{qa} - \cos qa \right] \quad (7.255)$$

where θ is defined indirectly through $q = |\mathbf{k} - \mathbf{k}'| = 2k \sin \frac{\theta}{2}$. Thus by scattering particles off a spherical well, such as a proton, we can establish the radius of the scatterer (proton).

In general, if $f(\mathbf{k}, \mathbf{k}')$ can be approximated by the first order Born amplitude $f^{(1)}$ we know that

1. $\frac{d\sigma}{d\Omega}$ of $f(\theta)$ is a function of q only; that is $f(\theta)$ depends on the energy $\hbar^2 k^2 / (2m)$ and the angle θ only through the combination $k^2(1 - \cos \theta)$.
2. $f(\theta)$ is always real
3. $\frac{d\sigma}{d\Omega}$ is independent of the sign of V .

7.7 IDENTICAL PARTICLES

Fundamental particles are indistinguishable; that is, you cannot label one electron as A and another as B and then follow each of their time-evolutions separately. There exists a permutation symmetry, which leads to some interesting physics. Consider the state $|k'\rangle|k''\rangle$ ($k' \neq k''$), that is, two quantum mechanical particles one with wavenumber k' and the other with k'' . Due to the permutation symmetry, we cannot distinguish between this state and the state $|k''\rangle|k'\rangle$. These two states are indistinguishable, and hence also degenerate: this is referred to as *exchange degeneracy*. In fact, all state of the form

$$\alpha|k'\rangle|k''\rangle + \beta|k''\rangle|k'\rangle, \quad |\alpha|^2 + |\beta|^2 = 1 \quad (7.256)$$

are degenerate. Define the permutation operator P_{12} which switches the two particles

$$P_{12}|k'\rangle|k''\rangle = P_{21}|k'\rangle|k''\rangle = |k''\rangle|k'\rangle \quad (7.257)$$

Suppose we have an operator A_1 which only operates on first ket, then we can show that

$$P_{12}A_1P_{12}^{-1} = A_2 \quad (7.258)$$

where A_2 is the same operator but operating on the second particle. Using the permutation operator we can define a symmetriser, S_{12} , and an antisymmetriser, A_{12} :

$$S_{12} = \frac{1}{2} (1 + P_{12}), \quad A_{12} = \frac{1}{2} (1 - P_{12}) \quad (7.259)$$

They take any state $|\psi'\rangle|\psi''\rangle$ and put them in a symmetric or antisymmetric configuration.

7.7.1 Symmetrisation Postulate

A system on N particles are either completely symmetric under a permutation of two particles (bosons) or completely antisymmetric under a permutation of two particles (fermions):

$$\begin{aligned} P_{ij}|\mathbf{N} \text{ identical bosons}\rangle &= +|\mathbf{N} \text{ identical bosons}\rangle \\ P_{ij}|\mathbf{N} \text{ identical fermions}\rangle &= -|\mathbf{N} \text{ identical fermions}\rangle \end{aligned} \quad (7.260)$$

It can be shown that the particle type is determined by its spin: fermions have half-integer spin and bosons have integer spin.

An immediate implication of Equation 7.260 is that two fermions can never occupy the same quantum state; if they *did* then the wavefunction would have to be equal to zero, implying that the probability of this occurring is zero.

For two particles fermions will always occupy the singlet state (antisymmetric under permutation) and boson will always occupy the three triplet states (symmetric under permutation). However, it should be noted that it is the *total* state that is to be totally symmetric or antisymmetric:

$$|\alpha\rangle = |\psi(x)\rangle \otimes |\chi\rangle \quad (7.261)$$

that is, for a fermionic system, the (tensor) product of the spatial wavefunction and the spin wavefunction have to be antisymmetric under permutation, thus if the spin configuration is in a triplet state, we know that the spatial wavefunction must be antisymmetric under permutation.

Helium Atom

Consider the Helium atom, which has a nucleus of charge $2e$ and two electrons. These electrons, like all electrons, are identical, therefore our rules about their permutation symmetry hold; their total wavefunction must be totally antisymmetric. Consider an excited state where one of the electrons is in its ground state and the other in an excited state

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} [\psi_{100}(\mathbf{x}_1)\psi_{n\ell m}(\mathbf{x}_2) \pm \psi_{n\ell m}(\mathbf{x}_1)\psi_{100}(\mathbf{x}_2)] \quad (7.262)$$

The energy of this state must be

$$E = E_{100} + E_{n\ell m} + \Delta E \quad (7.263)$$

where ΔE is due to the interaction between the electrons. First order *degenerate* perturbation theory tells us that ΔE is given by

$$\Delta E = \left\langle \frac{e^2}{r_{12}} \right\rangle = I \pm J \quad (7.264)$$

where the direct integral and exchange integral are

$$\begin{aligned} I &\equiv \iint d\mathbf{x}_1 d\mathbf{x}_2 |\psi_{100}(\mathbf{x}_1)|^2 |\psi_{n\ell m}(\mathbf{x}_2)|^2 \frac{e^2}{r_{12}} \\ J &\equiv \iint d\mathbf{x}_1 d\mathbf{x}_2 \psi_{100}(\mathbf{x}_1)\psi_{n\ell m}(\mathbf{x}_2) \frac{e^2}{r_{12}} \psi_{100}^*(\mathbf{x}_2)\psi_{n\ell m}^*(\mathbf{x}_1) \end{aligned} \quad (7.265)$$

where $r_{12} \equiv |\mathbf{x}_1 - \mathbf{x}_2|$

Multiparticle Systems

For higher number of particles things become quite complicated, however it is quite easy to find the totally antisymmetric state, which quite often is the lowest energy level. Using the *Slater Determinant*, for three particles we have

$$|k'k''k'''\rangle_{\text{anti-symmetric}} = \begin{vmatrix} |k'\rangle & |k''\rangle & |k'''\rangle \\ |k'\rangle & |k''\rangle & |k'''\rangle \\ |k'\rangle & |k''\rangle & |k'''\rangle \end{vmatrix} \quad (7.266)$$

Not that this is not zero; because we cannot let the kets commute with each other, $|k'\rangle|k''\rangle \neq |k''\rangle|k'\rangle$. Hence it is important that when we calculate the determinant we iterate of the first row, not the first column.

7.7.2 Quantum Fields

One of the major attractions of Quantum Field Theory is that it can deal with (special) relativistic quantum mechanics. However, the techniques, such as second quantisation, are useful even in the non-relativistic limit.

Second Quantisation

The space used to describe second quantisation is *Fock Space*, where we denote state by the number of particles in each state. So the ket $|n_1, n_2, \dots, n_i, \dots\rangle$ has n_1 electrons in the state $|k_1\rangle$, n_2 in the state $|k_2\rangle$ etc. There are two special states; the vacuum state

$$|0\rangle \equiv |0, 0, \dots, 0, \dots\rangle \quad (7.267)$$

which is void of any particles, and the single particle state

$$|k_i\rangle \equiv |0, 0, \dots, 0, n_i = 1, 0, \dots\rangle \quad (7.268)$$

Next we define the annihilation operator, a_i , and its hermitian conjugate, the creation operator a_i^\dagger . These annihilate and create a particle in state $|k_i\rangle$, respectively, hence

$$a_i^\dagger |n_1, n_2, \dots, n_i, \dots\rangle \propto |n_1, n_2, \dots, n_i + 1, \dots\rangle \quad (7.269)$$

similarly

$$a_i |n_1, n_2, \dots, n_i, \dots\rangle \propto |n_1, n_2, \dots, n_i - 1, \dots\rangle \quad (7.270)$$

Consider now the state $a_i^\dagger a_j^\dagger |0\rangle$ and $a_j^\dagger a_i^\dagger |0\rangle$. Effectively these are the same states, except that we've swapped k_i and k_j . Therefore we expect that

$$a_i^\dagger a_j^\dagger |0\rangle = \pm a_j^\dagger a_i^\dagger |0\rangle \quad (7.271)$$

depending on which type of particle we are describing. This tells us the depending on whether we are dealing with boson or fermions we either have a commutation relation, or an anti-commutation relation for the creation and annihilation operators:

$$\{a_i^\dagger, a_j^\dagger\} = 0 \quad \text{Fermions} \quad (7.272)$$

$$[a_i^\dagger, a_j^\dagger] = 0 \quad \text{Bosons} \quad (7.273)$$

we can also take the hermitian adjoint to get the same (anti-)commutation relation for the annihilation operators. Additionally we have that

$$\{a_i, a_j^\dagger\} = \delta_{ij} \quad \text{Fermions} \quad (7.274)$$

$$[a_i, a_j^\dagger] = \delta_{ij} \quad \text{Bosons} \quad (7.275)$$

Finally, for both bosons and fermions we define the number operator:

$$N = \sum_i a_i^\dagger a_i \quad (7.276)$$

which counts the total number of particles. Any operator that acts on each of the particles individually can be decomposed into a linear combination of projections:

$$\mathcal{K} = \sum_{n,m} a_m^\dagger a_n \langle k_m | K | k_n \rangle \quad (7.277)$$

The momentum and kinetic energy operators are two examples of many operators that can be written in this way. This, however, does not include inter-particle interactions, consider a real matrix, V , whose ij^{th} component specifies the interaction energy between $|k_i\rangle$ and $|k_j\rangle$, then the second quantisation version of this operator is

$$\mathcal{U} = \frac{1}{2} \sum_{i \neq j} V_{ij} N_i N_j + \frac{1}{2} \sum_i V_{ii} N_i (N_i - 1) \quad (7.278)$$

for $V_{ij} \in \mathbb{R}$, which ensures \mathcal{U} 's hermiticity. The second term accounts for "self-interaction". However, by writing the N operators out and using the (anti-)commutation relations it can be shown that this second-quantisation operator takes on a simpler form

$$\mathcal{U} = \frac{1}{2} \sum_{ij} V_{ij} a_i^\dagger a_j^\dagger a_j a_i \quad (7.279)$$

the ordering of the creation and annihilation operators is referred to as "normal ordering"; we annihilate the i^{th} particle first and then the j^{th} , and then create them in the opposite order. For fermions there are no diagonal contributions because we cannot annihilate more than one of each particle.

Degenerate Electron Gas

The electrons in a degenerate electron gas interact with one another (H_{el}) but also with the background energy level ($H_{\text{b-el}}$) which in itself has energy (H_{b}). We can show that the background interaction and the background energy sum up to a simple term, and we continue to write the electron-interaction part as is

$$H = -\frac{1}{2} \frac{e^2 N^2}{V \mu^2} + \sum_i \left(\frac{|\mathbf{p}_i|^2}{2m} + \frac{e^2}{2} \sum_{j \neq i} \frac{e^{-\mu |\mathbf{x}_i - \mathbf{x}_j|}}{|\mathbf{x}_i - \mathbf{x}_j|} \right) \quad (7.280)$$

where μ is the *screening* parametre which we will let $\mu \rightarrow 0$ once we have our solution. We can begin by converting the kinetic energy into a second quantisation operator: the kinetic energy is the $\frac{\hbar^2 |\mathbf{k}|^2}{2m}$ times the number of electrons with that momentum \mathbf{k} , summed over all possible values of \mathbf{k} :

$$\sum_i \frac{|\mathbf{p}_i|^2}{2m} \rightarrow \mathcal{K} = \sum_{\mathbf{k}\sigma} \frac{\hbar^2 |\mathbf{k}|^2}{2m} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} \quad (7.281)$$

where $\sigma \in \{+, -\}$ denotes the spin of the electron. The potential term (Coulomb interaction between the electrons) can be written as

$$\frac{e^2}{2V} \sum_{\mathbf{k}, \mathbf{p}, (\mathbf{q} \neq 0)} \sum_{\sigma_1 \sigma_2} \frac{4\pi}{q^2} a_{\mathbf{k}+\mathbf{q}, \sigma_2}^\dagger a_{\mathbf{p}-\mathbf{q}, \sigma_2}^\dagger a_{\mathbf{p}, \sigma_2} a_{\mathbf{k}, \sigma_1} \quad (7.282)$$

This can be approximated using perturbation theory giving a ground state energy of

$$\frac{E}{N} = \frac{e^2}{2a_0} \left(\frac{9\pi}{4} \right)^{\frac{2}{3}} \left(\frac{3}{5} \left(\frac{1}{r_s} \right)^2 - \frac{3}{2\pi} \frac{1}{r_s} \right) \quad (7.283)$$

where r_s is the Wigner-Seitz radius, which is a dimensionless distance scale defined as the ratio between the lattice constant and the Bohr radius.

7.7.3 Quantisation of the Electromagnetic Field

We can express the classical energy in an electromagnetic field in terms of two *numbers* $A_{\mathbf{k}, \sigma}$ and $A_{\mathbf{k}, \sigma}^*$:

$$\mathcal{E} = \frac{1}{4\pi} V \sum_{\mathbf{k}, \sigma} \frac{\omega_{\mathbf{k}}^2}{c^2} [A_{\mathbf{k}, \sigma}^* A_{\mathbf{k}, \sigma} + A_{\mathbf{k}, \sigma} A_{\mathbf{k}, \sigma}^*] \quad (7.284)$$

in the Quantum mechanical version thereof they become creation and annihilation operators, and we write this as

$$\mathcal{H} = \sum_{\mathbf{k}, \sigma} \hbar \omega_{\mathbf{k}} a_{\sigma}^\dagger(\mathbf{k}) a_{\sigma}(\mathbf{k}) + E_0 \quad (7.285)$$

photons are spin-1 particles, therefore they are bosons. The zero point energy $E_0 = \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}}$ is also referred to as the *vacuum energy* and is responsible for the Casimir Effect, where two parallel plates that are put very close to each other feel an attractive force, even though they are both electrically neutral.

7.8 RELATIVISTIC QUANTUM MECHANICS

In the following we will use natural units; $c = \hbar = 1$.

In special relativity we defined the total energy of a particle as

$$E = \sqrt{p^2 + m^2} \quad (7.286)$$

which in the low velocity limit became the rest energy plus the kinetic energy. However, for large values of p we need to include higher terms, or resort to a different method of calculating the energy.

7.8.1 Klein-Gordon Equation

The Klein-Gordon Equation tackles relativistic quantum mechanics by expressing the differential equation that describes time-evolution as a *second-order* differential equation:

$$[\partial_\mu \partial^\mu - m^2] \Psi(\mathbf{x}, t) = 0 \quad (7.287)$$

I use the $(-, +, +, +)$ signature where $\partial^\mu = \eta_{\mu\nu} \partial_\nu$ and $\eta_{\mu\nu}$ is the Minkowski metric. This differential equation is solved by

$$\Psi(\mathbf{x}, t) = N e^{-ip_\mu x^\mu} \quad (7.288)$$

if

$$p_\mu x^\mu + m^2 = -E^2 + p^2 + m^2 \quad (7.289)$$

so p^μ is the relativistic momentum

$$p^\mu = (E, p^1, p^2, p^3) \quad (7.290)$$

This leads to a probability four-current, j^μ :

$$j^\mu = \frac{i}{2m} [\Psi^* \partial^\mu \Psi - (\partial^\mu \Psi^*) \Psi] \quad (7.291)$$

so, for instance

$$\rho(\mathbf{x}, t) = \frac{i}{2m} [\Psi^* \partial_t \Psi - (\partial_t \Psi^*) \Psi] \quad (7.292)$$

though this is a conserved quantity, it is not positive definite. This is due the issue that the Klein-Gordon Equation predicts *negative*-energy states. This issue was quite serious, because it suggested that we cannot think of ρ as a *probability* densities because probabilities are positive definite.

7.8.2 Dirac Equation

Let us instead factorise the Klein Gordon Equation, so that the differential equation only depends on first order temporal and spatial derivatives:

$$E = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m, \quad \text{s.t.} \quad E^2 = \mathbf{p}^2 + m^2 \quad (7.293)$$

writing this out

$$E^2 = \boldsymbol{\alpha}^2 \mathbf{p}^2 + \beta^2 m^2 + m (\boldsymbol{\alpha} \beta \cdot \mathbf{p} + \beta \boldsymbol{\alpha} \cdot \mathbf{p}) \quad (7.294)$$

clearly

$$\alpha_i^2 = \mathbb{1}, \quad \beta^2 = \mathbb{1}, \quad \{\alpha_i, \beta\} = 0 \quad (7.295)$$

However, this is usually written slightly different, define $\gamma^0 \gamma^i = \alpha_i$ and $\beta = \gamma^0$, then this becomes

$$\{\gamma^\mu, \gamma^\nu\} = -\eta^{\mu\nu} \quad (7.296)$$

the minus is because I use the $(-, +, +, +)$ signature. This is the Clifford Algebra. The lower number of dimensions required for this is 4 in which we have

$$\boldsymbol{\alpha} = \sigma_x \otimes \sigma, \quad \beta = \sigma_z \otimes \mathbb{1} \quad (7.297)$$

Free Particle

Consider now a free particle of mass m and momentum $\mathbf{p} = p\hat{\mathbf{z}}$. In this case

$$\begin{pmatrix} m & 0 & p & 0 \\ 0 & m & 0 & -p \\ p & 0 & -m & 0 \\ 0 & -p & 0 & -m \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = E \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \quad (7.298)$$

The components u_1 and u_2 are the spinor components and u_3 and u_4 are the time-reversed (anti-particle) versions thereof. We see this because the upper two entries have positive energy levels, whereas the lower two have negative energy levels. We can find the eigenvectors quite easily:

$$u_R^+ = \begin{pmatrix} 1 \\ 0 \\ \frac{p}{E+m} \\ 0 \end{pmatrix}, \quad u_L^+ = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -\frac{p}{E+m} \end{pmatrix} \quad \text{for } E > 0 \quad (7.299)$$

the subscript stands for the handedness (helicity) of the of the state and the superscript tells us that these are the positive energy states. The negative energy states are quite similar:

$$u_R^- = \begin{pmatrix} \frac{-p}{E+m} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_L^- = \begin{pmatrix} 0 \\ \frac{p}{E+m} \\ 0 \\ 1 \end{pmatrix} \quad \text{for } E < 0 \quad (7.300)$$

Electromagnetic Interactions in the non-relativistic limit

Due to the presence of an electromagnetic field we need to use the kinetic momentum:

$$\mathbf{p} \rightsquigarrow \mathbf{p} - e\mathbf{A} \quad (7.301)$$

The Dirac Equation then becomes

$$\begin{pmatrix} m & \boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & -m \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = E \begin{pmatrix} u \\ v \end{pmatrix} \quad (7.302)$$

In the non-relativistic limit we can approximate the lower equation as

$$\boldsymbol{\sigma} \cdot \mathbf{p} u \approx 2mv \quad (7.303)$$

which means that the upper equation becomes

$$\frac{(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p})}{2m} u = \left(\frac{\mathbf{p}^2}{2m} + \frac{i\boldsymbol{\sigma}}{2m} \cdot (\mathbf{p} \times \mathbf{p}) \right) u \quad (7.304)$$

note that the cross product is nonzero, because $\nabla \times \mathbf{A} \neq \mathbf{A} \times \nabla$. Writing it all out we get that

$$\left[\frac{\mathbf{p}^2}{2m} - \boldsymbol{\mu} \cdot \mathbf{B} \right] u = Ku \quad (7.305)$$

where

$$\mu = \frac{ge}{2m} \frac{\hbar}{2} \sigma \quad (7.306)$$

thus the existence of spin *falls out* of the Dirac Equation; it is not something we need to justify empirically.

ELEMENTARY PARTICLE PHYSICS

8.1 QUANTISATION OF THE ELECTROMAGNETIC FIELD

The ultraviolet catastrophe as well as the photo-electric effect motivated physicists to try and rethink the theory of electromagnetism. Physicists realised that the Fourier decomposition of a vector potential that satisfies Maxwell's equations:

$$\mathbf{A}(\mathbf{x}, t) = \sum_{\lambda} \int \frac{d^3\mathbf{k}}{(2\pi)^3} (\mathbf{a}_{\mathbf{k},\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + \mathbf{a}_{\mathbf{k},\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{x}}) \quad (8.1)$$

can be expressed as a Harmonic Oscillator, by introducing the following variables:

$$\mathbf{Q}_{\mathbf{k}} = \mathbf{a}_{\mathbf{k},\lambda} + \mathbf{a}_{\mathbf{k},\lambda}^* \rightsquigarrow \dot{\mathbf{Q}}_{\mathbf{k}} = \mathbf{P}_{\mathbf{k}} \quad (8.2)$$

$$\mathbf{P}_{\mathbf{k}} = -i\omega_{\mathbf{k}} (\mathbf{a}_{\mathbf{k},\lambda} - \mathbf{a}_{\mathbf{k},\lambda}^*) \rightsquigarrow \dot{\mathbf{Q}}_{\mathbf{k}} = -\omega_{\mathbf{k}} \mathbf{Q}_{\mathbf{k}} \quad (8.3)$$

where we've used that $\dot{\mathbf{a}}_{\mathbf{k}} = i\omega_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}$. This means that the coefficients in the vector potential above can be expressed as a Harmonic Oscillator Hamiltonian:

$$H = \frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} (\mathbf{P}_{\mathbf{k}}^2 + \omega_{\mathbf{k}}^2 \mathbf{Q}_{\mathbf{k}}^2) \quad (8.4)$$

which in fact can be written as

$$H = \frac{1}{2} \int d^3\mathbf{x} (\mathbf{E}^2 + \mathbf{B}^2) \quad (8.5)$$

which we know to be the total energy of the electromagnetic field. We know how to quantise Harmonic Oscillators, we transform from Poisson Brackets to commutators:

$$[\hat{\mathbf{Q}}_{\mathbf{k}}, \hat{\mathbf{P}}_{\mathbf{k}'}] = i\hbar(2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}') \quad (8.6)$$

and the Hamiltonian becomes an operator. As we did for the regular harmonic oscillator, this can be expressed in terms of bosonic creation and annihilation operators:

$$\hat{H} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \hbar\omega_{\mathbf{k}} \sum_{\lambda=1,2} \left(\hat{\mathbf{a}}_{\mathbf{k},\lambda}^{\dagger} \hat{\mathbf{a}}_{\mathbf{k},\lambda} + \frac{1}{2} \right) \quad (8.7)$$

whence we can define the vector potential operator

$$\hat{\mathbf{A}}(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^3 \sqrt{2\omega_{\mathbf{k}}}} (\hat{\mathbf{a}}_{\mathbf{k},\lambda}(t) e^{i\mathbf{k}\cdot\mathbf{x}} + \hat{\mathbf{a}}_{\mathbf{k},\lambda}^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{x}}) \quad (8.8)$$

Clearly the quantum of light, a photon, is a boson, therefore we describe the many body state by a bosonic Fock state, where occupation numbers can be any non-negative integer. The creation operators, $\hat{\mathbf{a}}_{\mathbf{k},\lambda}^{\dagger}$ creates a photon with wave-vector \mathbf{k} and polarisation λ , whereas the annihilation operator annihilates the same photon.

8.2 DIRAC EQUATION

The Schrödinger Equation describes the dynamics of non-relativistic quantum mechanical particles, but at the time there was no way of describing relativistic quantum mechanical particles. The only equation one had that came close was the Klein-Gordon Equation

$$(\square + m^2)\phi = 0, \quad \square = \eta^{\mu\nu} \partial_\mu \partial_\nu \quad (8.9)$$

However, the issue with the Klein-Gordon equation was not that it predicted negative energy states, but that superposition states that included both positive and negative energy states did not have probability conservation. This is an issue because the leading interpretation at the time was the probabilistic interpretation of quantum mechanics.

Dirac's trick to "take the square root of the Klein-Gordon Equation" was that he needed to use matrices as coefficients. Consider

$$\begin{aligned} (\alpha_i p_i + \beta m)^2 &= p_i^2 + m^2 \\ &= \beta^2 m^2 + \sum_i \alpha_i^2 p_i^2 + \{\alpha_i, \beta\} p_i m + \sum_j \{\alpha_i, \alpha_j\} p_i p_j \end{aligned} \quad (8.10)$$

hence we require that $\beta^2 = \mathbb{1}$, $\{\alpha_i, \alpha_j\} = 2\delta_{ij}\mathbb{1}$ and $\{\alpha_i, \beta\} = \{\alpha_i, \alpha_j\} = 0$. By multiplying by β from the left we obtain the γ -matrices:

$$\gamma^0 = \beta, \quad \gamma^i = \beta \alpha^i \quad (8.11)$$

with these matrices we see that the matrices required for the Dirac Equation satisfy the Clifford Algebra:

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1} \quad (8.12)$$

It can be shown that α_i, β as well as γ^μ must be 4×4 matrices if we require $(3 + 1)$ -spacetime dimensions. Additionally we have that

$$\alpha^i = \sigma^x \otimes \sigma^i, \quad \beta = \sigma_z \otimes \mathbb{1} \quad (8.13)$$

(and corresponding relations for the γ -matrices). The fact that the Pauli matrices appear is not a coincidence, in fact if we were to look at the non-relativistic limit of the Dirac equation the 4-component differential equation would split into two 2-component differential equations: the first one is the Pauli Equation, given below, and the second one is approximately zero:

$$H = \frac{1}{2} ((\mathbf{p} - q\mathbf{A})^2 - q\hbar\boldsymbol{\sigma} \cdot \mathbf{B}) + q\phi \quad (8.14)$$

we see that the Dirac Equation *inherently* describes spin, and its interaction with the electromagnetic field. The Dirac equation is

$$H_{\text{Dirac}} = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m \quad (8.15)$$

but the Dirac Equation is often written as

$$(\gamma^\mu p_\mu - m)\psi = 0 \quad (8.16)$$

8.2.1 Useful γ -matrix identities

The relation between the α and β -matrices and the γ -matrices is

$$\gamma^0 = \beta, \quad \gamma^i = \beta\alpha^i \quad (8.17)$$

We often need to calculate traces of different combinations of γ -matrices

$$\text{tr} [\gamma^\mu] = 0 \quad (8.18a)$$

$$\text{tr} [\gamma_5] = 0 \quad (8.18b)$$

$$\text{tr} [\gamma_5 \gamma^\mu] = 0 \quad (8.18c)$$

$$\text{tr} [\gamma^\mu \gamma^\nu] = 4\eta^{\mu\nu} \quad (8.18d)$$

$$\text{tr} [\gamma_5 \gamma^\mu \gamma^\nu] = 0 \quad (8.18e)$$

$$\text{tr} [\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] = 4(\eta^{\mu\nu}\eta^{\rho\sigma} - \eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho}) \quad (8.18f)$$

$$\text{tr} [\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma_5] = -4i\epsilon^{\mu\nu\rho\sigma} \quad (8.18g)$$

$$\overline{\gamma^\mu} = \gamma^\mu \quad (8.18h)$$

The “fifth” γ -matrix is defined as

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (8.19)$$

and it behaves very much like other γ -matrices; it squares to the identity and anti-commutes with the other γ -matrices. Due to the fact that γ_5 is an involutory matrix, it has eigenvalues ± 1 . It turns out that these eigenvalues represent the *chirality* of the left and right spinors:

$$\gamma_5\psi_L = -\psi_L, \quad \gamma_5\psi_R = \psi_R \quad (8.20)$$

All 4×4 matrices can be written as a linear combination of the following set of matrices

$\{\mathbb{1}, \gamma^\mu, \gamma_5, \gamma^\mu\gamma_5, [\gamma^\mu, \gamma^\nu]\}$, *i.e.*

$$M = a_0\mathbb{1} + a_\mu\gamma^\mu + a_5\gamma_5 + a_{\mu 5}\gamma^\mu\gamma_5 + a_{\mu\nu}[\gamma^\mu, \gamma^\nu] \quad (8.21)$$

We can find the coefficients by means of

$$\text{tr} [M] = 4a_0, \quad \text{tr} [M\gamma^\mu] = 4a_\mu, \quad \text{tr} [M\gamma_5] = 4a_5, \quad \text{tr} [M\gamma^\nu\gamma_5] = -16a_{\nu 5} \quad (8.22)$$

and similar identities for the remaining matrices in the basis.

We will also have to use

$$\frac{1}{2}\text{tr} [\gamma^0\gamma_\mu\gamma^0\gamma_\nu] p^\mu q^\nu = 2[E_p E_q + (\mathbf{p} \cdot \mathbf{q})] \quad (8.23)$$

8.2.2 Dirac Conjugation

It turns out that the bilinear we are used to using $\psi^\dagger\psi$ is not Lorentz invariant, therefore we introduce the Dirac conjugate $\bar{\psi}$:

$$\bar{\psi} \equiv \psi^\dagger\gamma^0 \quad (8.24)$$

Its main property is that

$$(U_{\frac{1}{2}}\psi)^\dagger\gamma^0 = \bar{\psi}U_{\frac{1}{2}}^{-1} \quad (8.25)$$

where we have defined the following unitary matrix

$$U_{\frac{1}{2}}(\omega) \equiv \exp\left(\frac{i}{2}\omega_{ij}S^{ij}\right) \quad (8.26)$$

Using the this new conjugation we can build scalars, vectors and tensors:

$$\begin{aligned} \bar{\psi}\psi &: \text{Scalar} \\ \bar{\psi}\gamma^\mu\psi &: \text{Four-vector} \\ \bar{\psi}[\gamma^\mu, \gamma^\nu]\psi &: \text{Anti-symmetric tensor} \\ \bar{\psi}\gamma_5\psi &: \text{Pseudoscalar} \\ \bar{\psi}\gamma_5\gamma^\mu\psi &: \text{Pseudovector} \end{aligned}$$

Dirac Conjugation for Matrices

Similarly we can define Dirac conjugation for matrices

$$\bar{X} = \gamma^0 X^\dagger \gamma^0 \quad (8.27)$$

for some matrix X . Some useful identities are

$$\bar{\gamma}^\mu = \gamma^\mu \quad (8.28)$$

and

$$(\bar{u}_1 X u_2)^* = \bar{u}_2 \bar{X} u_1 \quad (8.29)$$

Slashed Quantities

To shorten our notation we define slashed quantities

$$\gamma^\mu \partial_\mu \rightarrow \not{\partial} \quad (8.30)$$

$$\gamma^\mu p_\mu \rightarrow \not{p} \quad (8.31)$$

$$(\gamma^\mu p_\mu)(\gamma^\nu p_\nu) \rightarrow (\not{p})^2 = p^2 \quad (8.32)$$

$$(i\gamma^\mu \partial_\mu - m) \rightarrow (i\not{\partial} - m) \quad (8.33)$$

8.2.3 *Spin & Lorentz Transformations*

Let us introduce the relativistic spin matrices

$$\Sigma_j = \frac{i}{4}\epsilon_{jkl}\gamma^k\gamma^l = \frac{i}{8}\epsilon_{jkl}[\gamma^k, \gamma^l] \quad (8.34)$$

We showed that

$$\mathbf{\Sigma}^2 = \frac{1}{2} \left(1 + \frac{1}{2} \right), \quad [\Sigma_j, \Sigma_k] = i\varepsilon_{jkl} \Sigma_l \quad (8.35)$$

which implies that the Dirac Equation describes spin- $\frac{1}{2}$ particles. The spin alone does not commute with the Dirac Hamiltonian, it is the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{\Sigma}$ that is a conserved quantity.

However, the Dirac spinor has four components as it describes a positive energy state, together with a negative energy state; an electron and a positron. However, even though we can write the spinor as a vector with four components:

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \quad (8.36)$$

it is *not* the case that spinors transform under Lorentz transformations in the same way as four-vectors do. Regular four-vector transform as follows:

$$x^\mu \rightarrow \Lambda^\mu_\nu x^\nu \quad (8.37)$$

however, spinors transform as spinors:

$$\psi \rightarrow \exp \left(\frac{i}{2} \omega_{\mu\nu} S^{\mu\nu} \right) \psi \quad (8.38)$$

where $\omega_{\mu\nu}$ contains all information about the Lorentz-angles, and

$$S^{\mu\nu} \equiv \frac{i}{4} [\gamma^\mu, \gamma^\nu] \quad (8.39)$$

However, we need the Dirac Equation to be Lorentz invariant, which means we need to establish how the rest of the Dirac Equation changes under Lorentz transforms, such that the entire equation is invariant. In the lectures we found this by simplifying this by only looking at the Weyl Equation.

8.2.4 Weyl Equation

The Weyl-Equation is the zero-mass limit of the Dirac Equation, where the Dirac Equation becomes block diagonal and it splits into two independent equations

$$i\sigma^\mu \partial_\mu \psi_R = 0, \quad i\bar{\sigma}^\mu \partial_\mu \psi_L = 0 \quad (8.40)$$

where

$$\sigma^\mu \equiv (1, \boldsymbol{\sigma}), \quad \bar{\sigma}^\mu \equiv (1, -\boldsymbol{\sigma}) \quad (8.41)$$

It turns out that there exists a group of matrices that describe the transformation of these Pauli vectors under Lorentz transformations

$$\sigma^\mu \rightarrow \Lambda^\mu_\nu \sigma^\nu = S \sigma^\mu S^\dagger \quad (8.42)$$

where $\det S = 1$. It turns out that $\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$ transforms like

$$\Psi \rightarrow \begin{pmatrix} S_L & 0 \\ 0 & S_R \end{pmatrix} \Psi \quad (8.43)$$

where $S_L(\beta, \phi) = S(\beta, \phi)$ and $S_R(\beta, \phi) = S(-\beta, \phi)$ and

$$S(d\beta, d\phi) = \mathbb{1} + ad\beta \cdot \sigma + ibd\phi \cdot \sigma, \quad a, b \in \mathbb{R} \quad (8.44)$$

for some infinitesimal rotation parametrised by $d\beta$ and $d\phi$. β parametrises the boost part and ϕ the rotational part of the transformation.

8.2.5 Parity Transformation

Under parity transformations we go from a left spinor to a right spinor. This also implies that S_R and S_L transform into each other under parity transformations. Equation 8.43 also holds in the massive case; The left and right spinors aren't mixed by Lorentz transformations.

8.2.6 Lorentz Algebra

The Lorentz algebra describes the generators of the Lorentz group, that is

$$\Lambda^\alpha{}_\beta \approx \delta^\alpha{}_\beta + \frac{i}{2} \omega_{\mu\nu} (M^{\mu\nu})^\alpha{}_\beta \quad (8.45)$$

where the generators $M^{\mu\nu}$ satisfy

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(\eta^{\nu\rho} M^{\mu\sigma} - \eta^{\mu\rho} M^{\nu\sigma} + \eta^{\mu\sigma} M^{\nu\rho} - \eta^{\nu\sigma} M^{\mu\rho}) \quad (8.46)$$

the generators are antisymmetric in their indices $M^{\mu\nu} = -M^{\nu\mu}$. There are 6 generators, but they split into two set of three generators; those with only spatial indices generate rotations (L_k), whereas those with a temporal index generate boosts (K_k).

$$M^{\mu\nu} = \begin{pmatrix} 0 & K_1 & K_2 & K_3 \\ -K_1 & 0 & L_3 & -L_2 \\ -K_2 & -L_3 & 0 & L_1 \\ -K_3 & L_2 & -L_1 & 0 \end{pmatrix} \quad (8.47)$$

$$L_k \equiv \frac{1}{2} \varepsilon_{ijk} M^{ij}, \quad K_i \equiv M^{i0}$$

The algebra has different representations, that can characterised by a pair of (half)-integers, (j, j') . The lowest of which are

$$(0, 0), \quad \left(\frac{1}{2}, 0\right), \quad \left(0, \frac{1}{2}\right), \quad \left(\frac{1}{2}, \frac{1}{2}\right) \quad (8.48)$$

For the spinor representation we denote the generators as $S^{\mu\nu}$, these are the same $S^{\mu\nu}$ as the ones defined previously.

8.2.7 *Electromagnetic Interactions*

Similarly to non-relativistic quantum mechanics the substitution from kinematic to canonical momentum $p_\mu \rightarrow p_\mu - eA_\mu$ accounts for electromagnetic interactions (at least if the right Gauge is chosen), hence

$$(p_\mu \gamma^\mu - eA_\mu \gamma^\mu - m)\psi = 0 \quad (8.49)$$

Written as a Hamiltonian

$$H_{\text{Dirac},A} = H_{\text{Dirac}} + eA_0 - e\boldsymbol{\alpha} \cdot \mathbf{A} \quad (8.50)$$

8.2.8 *Solving the Free Dirac Equation, Rest-Frame*

In the rest-frame¹ the Dirac Equation just becomes

$$\begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \begin{pmatrix} u_L \\ u_R \end{pmatrix} = E \begin{pmatrix} u_L \\ u_R \end{pmatrix} \quad (8.51)$$

There are two solutions

$$E = m, \quad u_L = u_R \quad (8.52)$$

$$E = -m, \quad u_L = -u_R \quad (8.53)$$

where the spinor $u_L = \xi$ is arbitrary. There are two linearly independent and orthogonal spinors ξ , the simplest example is

$$\xi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \xi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (8.54)$$

So the positive energy solutions are

$$\psi_s^+(\mathbf{x}, t) = e^{-imt} \begin{pmatrix} \xi_s \\ \xi_s \end{pmatrix}, \quad s = 1, 2 \quad (8.55)$$

and the negative energy solutions are

$$\psi_s^-(\mathbf{x}, t) = e^{imt} \begin{pmatrix} \eta_s \\ -\eta_s \end{pmatrix}, \quad s = 1, 2 \quad (8.56)$$

generally ξ_s and η_s are two sets of linearly independent spinors. Using the z-component of the spin-operator defined in Equation 8.34 we see that

$$\Sigma_z \psi_s^\pm = \begin{cases} +\frac{1}{2} \psi_s^\pm, & s = 1 \\ -\frac{1}{2} \psi_s^\pm, & s = 2 \end{cases} \quad (8.57)$$

hence for a general ξ_s the state describes a linear combination of up and down spin particles (anti-particles).

¹ For a specific representation of the Clifford Algebra

8.2.9 Spinors; Positive and Negative Modes

A general plane wave solution to the free Dirac equation is:

$$(i\partial - m)\psi(p)e^{-ip \cdot x} = e^{-ip \cdot x}(\not{p} - m)\psi(p) = 0 \quad (8.58)$$

which only has solutions if the particle is on-shell, *i.e* if

$$\det(\not{p} - m) = (p^2 - m^2)^2 = (E^2 - \mathbf{p}^2 - m^2)^2 \quad (8.59)$$

so

$$E = \pm \sqrt{\mathbf{p}^2 + m^2} \quad (8.60)$$

hence we can decompose the planar wave solution into the positive and negative modes

$$\psi(x) = u(\mathbf{p})e^{-ip \cdot x} + v(\mathbf{p})e^{+ip \cdot x}, \quad p_\mu = (\sqrt{\mathbf{p}^2 + m^2}, \mathbf{p}) \quad (8.61)$$

applying the Dirac Hamiltonian

$$\begin{aligned} (i\partial - m)\psi(x) \\ = e^{-ip \cdot x}(\not{p} - m)u(\mathbf{p}) - e^{+ip \cdot x}(\not{p} + m)v(\mathbf{p}) = 0 \end{aligned} \quad (8.62)$$

due to the fact that the two exponentials are linearly independent we know that their coefficients must both be equal to zero, hence

$$(\not{p} - m)u(\mathbf{p}) = 0 \quad (8.63)$$

$$(\not{p} + m)v(\mathbf{p}) = 0 \quad (8.64)$$

which also holds independently for each spin, s .

$$(\not{p} - m)u_s = 0, \quad \bar{u}_s(\not{p} - m) = 0 \quad (8.65)$$

$$(\not{p} + m)v_s = 0, \quad \bar{v}_s(\not{p} + m) = 0 \quad (8.66)$$

Orthonormality

The spinors u and v must be orthogonal and normalised:

$$\bar{u}_s u_{s'} = 2m\delta_{ss'}, \quad \bar{v}_s v_{s'} = -2m\delta_{ss'} \quad (8.67)$$

A similar identity holds when we are dealing with massless particles

$$u_s^\dagger(p)u_{s'}(p) = 2E(p)\delta_{ss'}, \quad v_s^\dagger(p)v_{s'}(p) = 2E(p)\delta_{ss'} \quad (8.68)$$

If you're missing a minus in Equation 8.68, that is because v has *negative* energy. Hence if we want to require that $\psi_s(x)$ still describes a probability amplitude, we must have that, for example

$$\psi_s(x) = \frac{1}{\sqrt{2EV}} u_s e^{-ip \cdot x} \quad (8.69)$$

where we've assumed that the spinor part u_s satisfies Equation 8.68.

8.2.10 Charge Conjugation

Observe that $v^* e^{-ip \cdot x}$ has *positive* energy, however, it is not the case that ψ^* satisfies the positive-energy Dirac equation. Instead we need to transform

$$\psi \rightarrow -i\gamma^2 \psi^* \equiv \psi^c \quad (8.70)$$

to go from a negative energy state to a positive energy state (and vice-versa). As the name hints towards, the charge conjugation transformation flips the sign of the charge of a particle, *i.e* if

$$(i\partial - e\mathcal{A} - m) \psi = 0 \implies (i\partial + e\mathcal{A} - m) \psi^c = 0 \quad (8.71)$$

8.3 PERTURBATION THEORY

8.3.1 Fermi's Golden Rule & Coulomb Scattering

A *central* quantity in Relativistic Perturbation Theory is the number of transitions per unit time, which can be calculated by means of Fermi's Golden Rule:

$$dw_{f \leftarrow i} = 2\pi |\langle f | \mathcal{U} | i \rangle|^2 \delta(E_f^{(0)} - E^{(0)} - \Omega) d\Pi_f \quad (8.72)$$

where \mathcal{U} is some perturbation and Ω is the driving frequency: we've assumed that \mathcal{U} is a harmonic perturbation with frequency Ω . We have also introduced the density of final states

$$d\Pi_f = g \frac{V d^3 \mathbf{k}_f}{(2\pi)^3} \quad (8.73)$$

where V is some volume and g is the internal degeneracy ($g = 2$ for spin- $\frac{1}{2}$ particles). Fermi's Golden Rule is often quoted in the continuous limit, where we take $\Omega \rightarrow 0$

$$dw_{f \leftarrow i} = 2\pi |\mathcal{U}_{fi}|^2 \delta(E_f^{(0)} - E^{(0)}) d\Pi_f \quad (8.74)$$

In the case of the Dirac Equation, an electromagnetic perturbation is given by

$$\mathcal{U} = e\gamma^0 \gamma^\mu A_\mu(x) \quad (8.75)$$

the simplest example is if we consider a static electric field;

$$\mathcal{U} = eA_0(x) \mathbb{1}, \quad A_0(x) = \frac{Ze}{4\pi |x|} \quad (8.76)$$

Our initial and final states are

$$\psi_i(x) = \sqrt{\frac{1}{2E_i V}} u_s(p_i) e^{-ip_i \cdot x}, \quad (8.77)$$

$$\bar{\psi}_f(x) = \sqrt{\frac{1}{2E_f V}} \bar{u}_r(p_f) e^{ip_f \cdot x} \quad (8.78)$$

where $p_i = (E_i, \mathbf{p}_i)$, $p_f = (E_f, \mathbf{p}_f)$. This gives us the following expression for the matrix element

$$\mathcal{U}_{fi} = \frac{Ze}{4\pi} \frac{1}{V} \sqrt{\frac{1}{4E_i E_f}} \bar{u}_f(p_f) \gamma^0 u_i(p_i) \frac{4\pi}{|\mathbf{q}|^2} \quad (8.79)$$

where $\mathbf{q} = \mathbf{p}_f - \mathbf{p}_i$ is the transferred momentum, and the last term came from the Fourier transform of the Coulomb potential. In reality we are interested in the squared matrix element

$$|\mathcal{U}_{fi}|^2 = \frac{1}{4E_i E_f V^2} |\bar{u}_f(p_f) \gamma^0 u_i(p_i)|^2 \left| \frac{4\pi Z\alpha}{\mathbf{q}^2} \right|^2 \quad (8.80)$$

where $\alpha = \frac{e^2}{\hbar c}$ is the fine-structure constant. We see that this interaction is of order α^2 .

Spin Sums & The Controversial Trace-Trick

In Equation 8.80 the initial spinor has spin s whereas the final has spin r , which gives us a matrix-element of the form

$$|\mathcal{M}|^2 = |\bar{u}_r \gamma^0 u_s|^2 \quad (8.81)$$

it is standard to average over initial sums and sum over final sums. We denote that we have performed this spin sum by setting a bar over the matrix element, *i.e*

$$\overline{|\mathcal{M}|^2} \equiv \frac{1}{2} \sum_s \sum_r |\bar{u}_r \gamma^0 u_s|^2 \quad (8.82)$$

Additionally, $\bar{u}_r \gamma^0 u_s \in \mathbb{C}$, which means that

$$\text{tr} \left[|\bar{u}_r \gamma^0 u_s|^2 \right] = |\bar{u}_r \gamma^0 u_s|^2 \quad (8.83)$$

Using the cyclicity of the trace we can instead write

$$|\bar{u}_r \gamma^0 u_s|^2 = \text{tr} \left[u_r \bar{u}_r \gamma^0 u_s \bar{u}_s \gamma^0 \right] \quad (8.84)$$

Note that $u_s \bar{u}_s$ is a matrix, which once we sum over spins becomes

$$\sum_s u_s(p) \bar{u}_s(p) = \not{p} + m, \quad \sum_s v_s(p) \bar{v}_s(p) = \not{p} - m \quad (8.85)$$

Thus the spinor part of Equation 8.80 becomes

$$\overline{|u_r(p_f) \gamma^0 u_s(p_i)|^2} = \frac{1}{2} \text{tr} \left[\gamma^0 (m + \not{p}_i) \gamma^0 (m + \not{p}_f) \right] \quad (8.86)$$

now we can use the γ -matrix trace identities from Equations 8.18 to obtain

$$\overline{|u_r \gamma^0 u_s|^2} = 4E_i^2 \left(1 - \frac{|\mathbf{p}_i|^2}{E_i^2} \sin^2 \left(\frac{\theta}{2} \right) \right) \quad (8.87)$$

where we've used energy conservation to set the two energies equal. To obtain this result we used Equation 8.23.

8.3.2 Cross-Section

The cross-section is defined as the number of particles scattered into solid angle $d\Omega$ divided by the total flux of incoming particles. The flux of a single particle is

$$\Phi = \frac{\bar{u}_s \gamma u_s}{2E_i V} = \frac{1}{V} \frac{\mathbf{p}_i}{E_i} = \frac{\mathbf{v}}{V} \quad (8.88)$$

The cross-section is then

$$d\sigma = \frac{dw_{f \leftarrow i}}{|\Phi|} \quad (8.89)$$

again looking at the Coulomb scattering example we get that

$$dw_{f \leftarrow i} = \frac{1}{V} \frac{4Z^2 \alpha^2}{\mathbf{q}^4} \left(1 - \frac{\mathbf{p}_i^2}{E_i^2} \sin^2 \left(\frac{\theta}{2} \right) \right) |\mathbf{p}_i| E_i d\Omega \quad (8.90)$$

and hence

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2 \alpha^2}{\mathbf{q}^4} \left(1 - \frac{\mathbf{p}_i^2}{E_i^2} \sin^2 \left(\frac{\theta}{2} \right) \right) \quad (8.91)$$

We can attempt to write this in a more general form, by recognising that a part of the expression came from the Fourier transform of the potential

$$\frac{d\sigma}{d\Omega} = \frac{E^2}{4\pi^2} |eA_0(\mathbf{q})|^2 \left(1 - \beta^2 \sin^2 \left(\frac{\theta}{2} \right) \right), \quad \beta \equiv \frac{|\mathbf{p}_i|}{E_i} \quad (8.92)$$

in the non-relativistic limit, $\beta \rightarrow 0$, this reduced to the classical Rutherford formula.

8.3.3 Scattering off an Arbitrary Potential

Consider now a general electromagnetic scattering process, where the matrix element becomes

$$\mathcal{U}_{fi} = \frac{1}{V} \sqrt{\frac{1}{4E_i E_f}} \bar{u}_r(p_f) \gamma^\mu u_s(p_i) \int d^3\mathbf{x} A_\mu(x) e^{ix \cdot (\mathbf{p}_i - \mathbf{p}_f)} \quad (8.93)$$

thus we are now interested in

$$\begin{aligned} \frac{1}{2} \sum_{rs} (\bar{u}_r \gamma^\mu u_s) (\bar{u}_r \gamma^\nu u_s)^* &= \frac{1}{2} \text{tr} \left[\gamma^\mu (\not{p}_i + m) \gamma^\nu (\not{p}_f + m) \right] \\ &= 2 \left[\eta^{\mu\nu} m^2 + p_i^\mu p_f^\nu + p_i^\nu p_f^\mu - \eta^{\mu\nu} (p_i \cdot p_f) \right] \end{aligned} \quad (8.94)$$

this gives us \mathcal{U}_{fi} and hence we can calculate $dw_{f \leftarrow i}$.

8.3.4 Electron-Muon Scattering

Unfortunately, in order to describe electron-muon scattering we need to make our methods slightly more sophisticated, this is because the matrix element

$$\langle \text{electron} | \text{interaction} | \text{muon} \rangle = 0 \quad (8.95)$$

always, irrespective of the interaction. That is, electrons and muons do not interact directly, their interaction must be mediated by a particle that they both interact with; the photon (for example). Our states are now multi-particle states, we must keep track of the number of electrons, muons and photons:

$$|\Psi\rangle \equiv |e^-\rangle \otimes |\gamma\rangle \otimes |\mu^-\rangle \quad (8.96)$$

the initial and final states are

$$|i\rangle = |e(p)\rangle \otimes |0\rangle \otimes |\mu(q)\rangle \quad (8.97)$$

$$|f\rangle = |e(p')\rangle \otimes |0\rangle \otimes |\mu(q')\rangle \quad (8.98)$$

however, due to the fact that the particles do not interact directly our interaction must involve an intermediate state

$$|n\rangle = |e(p_n)\rangle \otimes |\gamma(k_n)\rangle \otimes |\mu(q_n)\rangle \quad (8.99)$$

Due to the intermediate state(s) we need to use the second order equivalent of Fermi's Golden Rule:

$$dw_{f \leftarrow i} = 2\pi \left| \sum_n \frac{\langle f|V|n\rangle \langle n|V|i\rangle}{E_i - E_n} \right|^2 \delta(E_i - E_f) d\Pi_f \quad (8.100)$$

the sum is over all possible intermediate states. The interacting Dirac Hamiltonian in this case is

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m - e\beta\gamma^\mu A_\mu(x) \quad (8.101)$$

where

$$A_\mu(x) = \sum_\lambda \int \frac{d^3\mathbf{k}}{(2\pi)^3 \sqrt{\omega_{\mathbf{k}}}} \left[\varepsilon_\mu^{(\lambda)} a_{\mathbf{k}} e^{-ik \cdot x} + \varepsilon_\mu^{(\lambda)*} a_{\mathbf{k}}^\dagger e^{+ik \cdot x} \right] \quad (8.102)$$

where the two polarisation vectors satisfy $\varepsilon_\mu^{(\lambda)} \cdot k^\mu = 0$ and $\varepsilon_\mu^{(\lambda)} \cdot (\varepsilon_{(\lambda')}^\mu)^* = -\delta_{\lambda\lambda'}$.

However, this holds both for the electron and the muon, so A_μ is the quantity that couples the otherwise independent particles. We will denote the γ -matrices for the electron as γ^μ and for the muon as Γ^μ . One should think of these as

$$\gamma^\mu = \gamma^\mu \otimes \mathbb{1}, \quad \Gamma^\mu = \mathbb{1} \otimes \gamma^\mu \quad (8.103)$$

i.e. that they are the same quantities but live in different Hilbert Spaces; however these Hilbert Spaces are mixed due to the interaction. Define the matrix element \mathcal{M}_1 as

$$\mathcal{M}_1 \equiv \sum_\lambda \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{V_{in} V_{nf}}{E_i - E_n} \quad (8.104)$$

which can be shown to be equal to

$$\mathcal{M}_1 = \frac{(2\pi)^3 \delta^{(3)}(\mathbf{p} + \mathbf{q} - \mathbf{p}' - \mathbf{q}')}{8\omega_{\mathbf{k}} V^2 \sqrt{E_e E_e' E_\mu E_\mu'}} e^{i(E_i - E_f)t} \sum_\lambda \varepsilon_\mu \varepsilon_\nu^* \frac{(\bar{u}_s \gamma^\mu u_{s'}) (\bar{U}_r \Gamma^\nu U_{r'})}{(E_e - E_e') - \omega_{\mathbf{k}}} \Big|_{\mathbf{k}=\mathbf{q}'-\mathbf{q}=\mathbf{p}-\mathbf{p}'}$$

we can ignore the $e^{i(E_i - E_f)t}$ due to the energy conservation delta function in Fermi's Golden Rule. We need to square this quantity and then use all of our tricks to simplify this. Another useful identity is that

$$\sum_{\text{polarisations}} \varepsilon_\mu \varepsilon_\nu^* \rightarrow -\eta^{\mu\nu} \quad (8.105)$$

this is not an equality, but we can replace the polarisation sum by the Minkowski metric and still obtain the same answers.

The term M_1 corresponds to the emission of a photon by the electron (and hence absorption by the muon), however the opposite process is also possible, so we define

$$M_2 = \frac{(2\pi)^3 \delta^{(3)}(\mathbf{p} + \mathbf{q} - \mathbf{p}' - \mathbf{q}')}{8\omega_{\mathbf{k}} V^2 \sqrt{E_e E'_e E_\mu E'_\mu}} e^{i(E_i - E_f)t} \sum_{\lambda} \varepsilon_{\mu} \varepsilon_{\nu}^* \frac{(\bar{u}_s \gamma^{\mu} u_{s'}) (\bar{U}_r \Gamma^{\nu} U_{r'})}{-(E_e - E'_e) - \omega_{\mathbf{k}}} \Big|_{\mathbf{k} = -(\mathbf{q}' - \mathbf{q}) = -(\mathbf{p} - \mathbf{p}')}$$

Due to the fact that $\omega_{\mathbf{k}} = |\mathbf{k}|$ the photon energies are (naturally) the same in the two processes, thus the total matrix element is

$$M = M_1 + M_2 = (2\pi)^3 \delta^{(3)}(\mathbf{p} + \mathbf{q} - \mathbf{p}' - \mathbf{q}') \frac{(\bar{u}_s \gamma^{\mu} u_{s'}) (\bar{U}_r \Gamma_{\mu} U_{r'})}{2V \sqrt{E_e E'_e} 2V \sqrt{E_{\mu} E'_{\mu}}} \frac{e^2}{(p - p')^2}$$

notice that the last product in the expression above is the Green's function of the wave equation. Now we need to square the matrix element and perform the spin sum². We obtain

$$|M|^2 = \frac{(2\pi)^6}{16V^4} \frac{I_{\mu\nu}^{(e)}}{E_e E'_e E_{\mu} E'_{\mu}} \frac{I_{(\mu)}^{\mu\nu}}{(p - p')^4} [\delta^{(3)}(\mathbf{p} + \mathbf{q} - \mathbf{p}' - \mathbf{q}')]^2 \quad (8.106)$$

where we have defined

$$I_{\mu\nu}^{(e)} \equiv \frac{1}{2} \text{tr} [\gamma_{\mu} (m + \not{p}) \gamma_{\nu} (m + \not{p}')] , \quad I_{(\mu)}^{\mu\nu} \equiv \frac{1}{2} \text{tr} [\Gamma^{\mu} (M + \not{q}) \Gamma^{\nu} (M + \not{q}')] \quad (8.107)$$

the squared delta-function gives us a factor V once we integrate. Putting this altogether we get hat

$$dw_{f \leftarrow i} = \frac{(2\pi)^4}{V} \frac{e^4 I_{\mu\nu}^{(e)} I_{(\mu)}^{\mu\nu}}{2E_e 2E_{\mu}} \frac{\delta^{(4)}(p + q - p' - q')}{(p - p')^4} d\Pi_{\text{LIPS}} \quad (8.108)$$

where

$$d\Pi_{\text{LIPS}} = \frac{d^3 \mathbf{p}'}{(2\pi)^3 (2E'_e)} \frac{d^3 \mathbf{q}'}{(2\pi)^3 (2E'_{\mu})} \quad (8.109)$$

Equation 8.108 is almost Lorentz invariant, it is only the volume and the $2E_e 2E_{\mu}$ that is non-Lorentz invariant. This issue is solved by looking at the cross-section instead; the flux in this case is

$$\Phi = \frac{1}{V} \left| \frac{\mathbf{p}}{E_e} - \frac{\mathbf{q}}{E_{\mu}} \right| \quad (8.110)$$

which can be written as

$$VE_1 E_2 \Phi = \sqrt{(p \cdot q)^2 - m^2 M^2} \quad (8.111)$$

In the limit where the electron is relativistic we can simplify the I terms:

$$I^2 = 8M^2 \left[(E_e - E'_e)^2 + 2E_e E'_e \cos^2 \frac{\theta}{2} \right] \quad (8.112)$$

putting this all together and integrating with respect to the muon momentum and energy we obtain

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_e^2 \sin^4 \frac{\theta}{2}} \frac{E'_e}{E_e} \left(\cos^2 \frac{\theta}{2} + \frac{Q^2}{2M^2} \sin^2 \frac{\theta}{2} \right) \quad (8.113)$$

where

$$Q^2 = 2ME_e \left(1 - \frac{M}{M + 2E_e \sin^2 \frac{\theta}{2}} \right) \quad (8.114)$$

2 here we have two initial spins and two final spins so the spin sum corresponds to $\frac{1}{4} \sum_{srs'r'}$

8.3.5 Feynman Diagrams

Feynman Diagrams are tools that we will use to simplify the process of obtaining an expression for the squared matrix element. The convenient thing is that complicated processes can be expressed pictorially, and that the complicated processes can then quite easily be converted to a mathematical expression.

For a $2 \rightarrow N \geq 2$ process we have that

$$d\sigma = \frac{(2\pi)^4 \delta^{(4)}(p_1 + p_2 - \sum_{i=1}^N p'_i)}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} \overline{|\mathcal{M}|^2} \prod_{i=1}^N \frac{d^3 \mathbf{p}'_i}{(2E'_i)(2\pi)^3} \quad (8.115)$$

all of these quantities are assumed to be known except the squared matrix element; Feynman Diagrams help us find this. Similarly if we want to calculate a decay rate for a $1 \rightarrow N \geq 2$ process we have

$$d\Gamma = \frac{(2\pi)^4}{2M} \delta^{(4)}\left(P - \sum_{i=1}^N p'_i\right) \overline{|\mathcal{M}|^2} \prod_{i=1}^N \frac{d^3 \mathbf{p}'_i}{(2E'_i)(2\pi)^3} \quad (8.116)$$

for an initial particle with mass M and $P = (M, 0, 0, 0)$ (at rest)

Feynman Rules for Quantum Electrodynamics

We need to distinguish between internal and external lines. External lines are lines that are only connected to the diagram at one of its two vertices. Internal lines are connected at both vertices and describe virtual particles.

1. a) Each incoming spinor is denoted by u_s, U_r, \dots
b) Each incoming photon is denoted by ε_μ, \dots
2. a) Each outgoing spinor is denoted by $\bar{u}_s, \bar{U}_r, \dots$
b) Each outgoing photon is denoted by ε_μ^*, \dots
3. At each vertex write $(-ie\gamma^\mu)$ for μ, ν, \dots
4. Start at the *end* of a fermion line and follow the lines backwards. Write whatever appears from left to right
5. Internal lines get *propagators*:

$$\text{photons : } \frac{-i\eta_{\mu\nu}}{q^2 + i\varepsilon}, \quad \text{fermions : } \frac{i(q + m)}{q^2 - m^2 + i\varepsilon} \quad (8.117)$$

where q is chosen such that momentum is conserved at each vertex and the photon indices match the γ^μ indices at the vertices from before

6. Set this all equal to $i\mathcal{M}$

Table 8.1: Feynman Rules: QED

8.4 FERMION THEORY

Prior to Fermi theory β -decay wasn't properly understood; one assumed that β -decay occurred due to a contact force between for instance an incoming neutron and the three products, a proton an electron and an anti-electron-neutrino.

Though this model had some promising predictions (that fit very well to data) there was a fundamental flaw with the model: the cross section did not conform to the unitarity (Froissart) bound which implies there was something fundamentally wrong with the way of thinking about what was later to become weak interactions. This will lead us to the prediction and discovery of the W -boson.

Before we describe this, let us talk about parity violation. All theory we have learnt thus far is invariant under parity transformations ($\mathbf{x} \rightarrow -\mathbf{x}$, $\mathbf{E} \rightarrow -\mathbf{E}$, $\mathbf{B} \rightarrow \mathbf{B}$, $\mathbf{p} \rightarrow -\mathbf{p}$, $\mathbf{S} \rightarrow -\mathbf{S}$ etc.). However, parity is violated by weak interactions, instead it is charge-parity that is a symmetry of weak interactions³. This tells us that left & right particles are distinct for weak interactions; in fact the weak force only couples to left particles.

By introducing a massive particle, the W -boson we can solve the unitarity issue, which adds the following to the Feynman Rules

1. c . Each incoming W -boson is denoted by ε_μ, \dots
2. c . Each outgoing W -boson is denoted by ε_μ^*, \dots
3. At each weak-interaction-vertex write

$$\frac{ig}{2\sqrt{2}}\gamma^\mu(1 - \gamma_5) \quad (8.118)$$

4. Same as before
5. Internal lines get *propagators*:

$$\begin{array}{ll} W\text{-bosons :} & -i \frac{\eta_{\mu\nu} - \frac{p_\mu p_\nu}{M_W^2}}{q^2 - M_W^2 + i\varepsilon}, \quad \text{neutrinos :} \quad i \frac{\not{q}}{q^2} \end{array} \quad (8.119)$$

where q is chosen such that momentum is conserved at each vertex and the photon indices match the γ^μ indices at the vertices from before

6. Same as before

Table 8.2: Feynman Rules: W -bosons

W -boson spin sums give us

$$\sum_\lambda \varepsilon_\mu \varepsilon_\nu^* = \left(-\eta_{\mu\nu} + \frac{k_\mu k_\nu}{M_W^2} \right) \quad (8.120)$$

³ It turns out that CP -symmetry also is violated

The coupling constant g is called the *weak coupling constant* and is related to the Fermi- g factor through

$$\frac{8G_F}{\sqrt{2}} = \frac{g^2}{M_W^2 c^4}, \quad G_F \approx 1.16 \times 10^{-5} \text{GeV}^{-2} \quad (8.121)$$

8.4.1 Mandelstam Variables

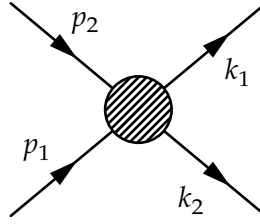


Figure 8.1: A generic $2 \rightarrow 2$ process

Consider a $2 \rightarrow 2$ process. It is useful to define

$$s = (p_1 + p_2)^2 \quad (8.122)$$

$$t = (p_1 - k_1)^2 \quad (8.123)$$

$$u = (p_1 - k_2)^2 \quad (8.124)$$

For instance if

$$(8.125)$$

we get⁴ that the momentum of the boson in the middle has momentum $k^2 = s$, which is why this is called the s -channel diagram. Similarly if

$$(8.126)$$

the boson has momentum $k^2 = t$ – this is the t -channel⁵. Finally

$$(8.127)$$

we have $k^2 = u$ – the u -channel. This is just like the t -channel except that the last two particles were swapped around.

⁴ In the relativistic limit $|\overline{\mathcal{M}}|^2 = 8e^4 \frac{t^2 + u^2}{s^2}$

⁵ In the relativistic limit $|\overline{\mathcal{M}}|^2 = 8e^4 \frac{s^2 + u^2}{t^2}$

8.4.2 Gauge Invariance and Gauge Fields

As we have seen in regular quantum mechanics, we would like our observables to be invariant under Gauge transformations of the form

$$\psi(x) \rightsquigarrow \psi(x)e^{ie\Gamma(x)} \quad (8.128)$$

which implies the presence of an extra gauge field which transforms as

$$\mathbf{A} \rightsquigarrow \mathbf{A} + \nabla\Gamma(x) \quad (8.129)$$

which means that we should replace partial derivatives to the *covariant derivative*, similarly to in General Relativity

$$\partial_\mu \rightsquigarrow D_\mu \equiv \partial_\mu - ieA_\mu \quad (8.130)$$

which implies that the Dirac Equation transforms as

$$\left[i\gamma^\mu \partial_\mu - m \right] \psi = 0 \rightsquigarrow \left[i\gamma^\mu (\partial_\mu - ieA_\mu) - m \right] \psi = 0 \quad (8.131)$$

This implies that simply because we require our physics to be invariant under gauge transformations we see that our particles interact with a new field, this is known as *minimal coupling* or *gauge interactions*.

It is natural to define the slashed covariant derivative, just like we did for other quantities

$$\not{D} \equiv \gamma^\mu (\partial_\mu - ieA_\mu(x)) \quad (8.132)$$

in which case the Gauge invariant Dirac Equation becomes

$$i\not{D}\psi - m\psi = 0 \quad (8.133)$$

the covariant derivative does not in general commute with itself,

$$[D_\mu, D_\nu] = -ieF_{\mu\nu} \quad (8.134)$$

which allows us to write the Klein-Gordon Equation for a spinor in a covariant form

$$(D_\mu D^\mu + m^2 - ieS^{\mu\nu}F_{\mu\nu})\psi = 0 \quad (8.135)$$

we can in fact do something similar for vector-bosons, such as the W -bosons. For them we have that

$$(W_\mu^+)^* = W_\mu^- \quad (8.136)$$

these vector fields transform as

$$W_\mu^\pm \rightsquigarrow e^{\pm ie\Gamma(x)} W_\mu^\pm \quad (8.137)$$

It turns out that the Klein-Gordon Equation for these vector bosons is

$$(D_\nu D^\nu + M_W^2)W_\mu^\pm - iekF_{\mu\nu}W_\pm^\nu \quad (8.138)$$

for some unknown dimensionless κ . We see that W -bosons also interact solely with photons (not only with fermions).

8.4.3 Electroweak Feynman Rules

1. The interaction $WW\gamma$ connects three vector fields $W_\mu^+(p), W_\lambda^-(k), A_\nu(q)$ thus the vertex should carry those three indices

2. We choose all momenta to be outgoing, thus

$$p_\mu + q_\mu + k_\mu = 0 \quad (8.139)$$

3. The resulting vertex has the form

$$V_{\lambda\mu\nu}(k, p, q) = ie \left[(k - p)_\nu \eta_{\mu\lambda} + (p_\lambda - \kappa q_\lambda) \eta_{\mu\nu} + (\kappa q_\mu - k_\mu) \eta_{\nu\lambda} \right] \quad (8.140)$$

Table 8.3: Feynman Rules: $W^+W^-\gamma$

Similarly for $WW\gamma\gamma$.

$$V_{\mu\nu\rho\sigma} = -e^2 (2\eta_{\mu\nu}\eta_{\rho\sigma} - \eta_{\mu\rho}\eta_{\nu\sigma} - \eta_{\mu\sigma}\eta_{\nu\rho}) \quad (8.141)$$

Table 8.4: Feynman Rules: $W^+W^-\gamma\gamma$

An example of the application of these rules is the process $e^+e^- \rightarrow W^+W^-$:

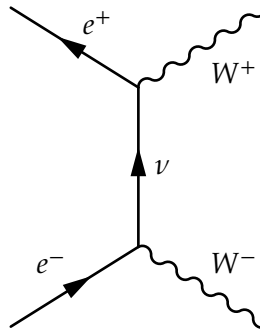


Figure 8.2: One of the Feynman Diagrams for the $e^+e^- \rightarrow W^+W^-$ process

The amplitude for which is given by

$$i\mathcal{M} = \frac{-g^2}{8} \bar{v}_s(p) \gamma^\mu (1 - \gamma_5) \frac{iq}{q^2} \gamma^\nu (1 - \gamma_5) u_r(k) \varepsilon_\mu^*(r_1) \varepsilon_\nu^*(r_2) \quad (8.142)$$

8.5 YANG-MILLS THEORY

Consider two distinct fermions, $\psi^{(1)}$ and $\psi^{(2)}$, with the same mass (or similar masses?). They live in each their Hilbert space and both satisfy a Dirac Equation:

$$(H_{\text{Dirac}} \oplus H_{\text{Dirac}}) \boldsymbol{\psi} = 0, \quad \boldsymbol{\psi} = \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \end{pmatrix} \quad (8.143)$$

$\boldsymbol{\psi}$ is referred to as an iso-vector. We definitely need to require the probability to be conserved:

$$P = \int d^3\mathbf{x} \boldsymbol{\psi}^\dagger \cdot \boldsymbol{\psi} \quad (8.144)$$

which works if we transform the iso-vector by means of a special-unitary ($SU(2)$) transformation. However, the question is whether this transformation maps solutions to the Dirac Equation to solutions. This is no different from what we did with Gauge Transformations, however the Gauge Transformation is a $U(1)$ transformation and this is a $SU(2)$ transformation. So is it the case that

$$(H_{\text{Dirac}} \oplus H_{\text{Dirac}}) U(x) \boldsymbol{\psi} = 0, \quad U(x) \in SU(2) \quad (8.145)$$

Let us denote $H_{\text{Dirac}} = (i\gamma^\mu \partial_\mu - m)$ (omitting the $\mathbb{1}_{2 \times 2} \otimes$). If we let $U(x) = \exp(i\boldsymbol{\alpha}(x) \cdot \mathbf{t})$ where $\mathbf{t} = \frac{\sigma}{2}$ then we get that

$$[i\partial - m] U(x) \boldsymbol{\psi} = U(x) [i\partial - \partial\alpha(x) - m] \boldsymbol{\psi} = 0 \quad (8.146)$$

which is reminiscent of what we obtained for Gauge transformations. We defined

$$\partial\alpha(x) = \gamma^\mu \partial_\mu \alpha(x) = \gamma^\mu \partial_\mu \boldsymbol{\alpha}(x) \cdot \mathbf{t} = -i\gamma^\mu U^\dagger \partial_\mu U \quad (8.147)$$

As we did previously we should now introduce a covariant derivative that accounts for this extra term. Define now

$$gB_\mu^{(j)} = -\partial_\mu \alpha^{(j)} \quad (8.148)$$

Then the modified Dirac Equation reads

$$(i\gamma^\mu \partial_\mu + g\gamma^\mu B_\mu - m) \boldsymbol{\psi} = 0, \quad B_\mu \equiv \mathbf{B}_\mu \cdot \mathbf{t} \quad (8.149)$$

Thus the covariant derivative should be

$$\mathbb{D}_\mu \equiv \partial_\mu \mathbb{1}_{2 \times 2} - igB_\mu \quad (8.150)$$

which means

$$[i\mathbb{D}_\mu - m] \boldsymbol{\psi} = 0 \quad (8.151)$$

is our new Dirac Equation. An important identity is that

$$\mathbb{D}'_\mu \boldsymbol{\psi}' = \mathbb{D}'_\mu U(x) \boldsymbol{\psi} = U(x) \mathbb{D}_\mu \boldsymbol{\psi} \quad (8.152)$$

where the transformed B_μ is

$$B'_\mu = UB_\mu U^\dagger - \frac{i}{g} (\partial_\mu U) U^\dagger \quad (8.153)$$

8.5.1 *Non-abelian Field Strength*

For electrodynamics we had Equation 8.134

$$F_{\mu\nu} \equiv \frac{i}{e}[D_\mu, D_\nu] \quad (8.154)$$

and similarly for our non-abelian fields we have

$$\mathbb{G}_{\mu\nu} = \frac{i}{g}[\mathbb{D}_\mu, \mathbb{D}_\nu] = \partial_\mu \mathbb{B}_\nu - \partial_\nu \mathbb{B}_\mu - ig[\mathbb{B}_\mu, \mathbb{B}_\nu] \quad (8.155)$$

The covariant derivative acting on the field gives us

$$\mathbb{D}_\lambda \mathbb{G}_{\mu\nu} = \partial_\lambda \mathbb{G}_{\mu\nu} - ig[\mathbb{B}_\lambda, \mathbb{G}_{\mu\nu}] \quad (8.156)$$

thus $\mathbb{G}_{\mu\nu}$ is not Gauge invariant which points towards the fact that these non-abelian fields are *charged*. If we were to define the Yang-Mills Lagrangian now we would get

$$\mathcal{L}_{\text{Yang-Mills}} = -\frac{1}{2} \text{tr} (\mathbb{G}_{\mu\nu} \mathbb{G}^{\mu\nu}) \quad (8.157)$$

which is the Lagrangian of a massless field, however, we observe only massive fields. We need the Higgs Mechanism to understand how these particles obtain Mass.

9.1 QUANTUM FIELDS

Like a classical field, a quantum field is a continuous function of position. However, in the language of second quantisation it is a field *operator*. This implies that it is subject to quantum fluctuations and that it in general does not commute with other operators.

The “first quantisation” corresponds to the transformation

$$E \rightsquigarrow i\hbar\partial_t, \quad p \rightsquigarrow \hat{p} = -i\hbar\partial_t \quad (9.1)$$

Additionally the classical Poisson Bracket is replaced by the commutator:

$$\{p_i, q_j\}_{\text{PB}} = \delta_{ij} \rightsquigarrow [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \quad (9.2)$$

The “second quantisation” is the next step we require to describe macroscopic systems, quasiparticles, dynamic response of quantum systems as well as collective behaviour and broken-symmetry phase transitions.

The classical Hamiltonian for a string is

$$H = \int dx \left\{ \frac{T}{2} (\nabla_x \phi)^2 + \frac{1}{2\rho} \pi^2 \right\} \quad (9.3)$$

which we can quantise by enforcing the following commutator

$$[\phi(x), \pi(y)] = i\hbar\delta(x - y) \quad (9.4)$$

hence the procedure of second quantisation is quite similar to what we did during our initial quantum mechanics courses, however, the algebra that we will come to do is quite different. In many-body physics the relevant operator is the quantum field operator $\psi(x)$, the *annihilation* operator whose hermitian conjugate is the *creation* operator, $\psi^\dagger(x)$.

The Pauli exclusion principle, which only holds for fermions, can be boiled down to the simple difference in how these field operators commute for the two different particle types:

$$[\psi(x), \psi^\dagger(y)] = \delta(x - y) \quad \text{Bosons} \quad (9.5)$$

$$\{\psi(x), \psi^\dagger(y)\} = \delta(x - y) \quad \text{Fermions} \quad (9.6)$$

The operators operate on a state within *Fock-Space*, which is the many-particle generalisation of Hilbert Space (it is a tensor product of Hilbert Spaces). The *Fock-Vacuum* is the state in which there are no particles, which is denoted $|0\rangle$. Upon this zero-particle state we can create particles and make many-particle states

$$|x_1, x_2, \dots, x_N\rangle = \psi^\dagger(x_N) \dots \psi^\dagger(x_1)|0\rangle \quad (9.7)$$

9.1.1 *Collective Modes*

Consider now a one-dimensional chain of particles, whose position is expressed by their equilibrium position plus some displacement $x_j = x_j^0 + \phi_j$. Let us now couple neighbouring sites by imposing the following Hamiltonian

$$\hat{H} = \sum_j \left[\frac{\pi_j^2}{2m} + \frac{m\omega^2}{2} (\phi_j - \phi_{j+1})^2 \right] = \sum_j \left[\frac{\pi_j^2}{2m} + \frac{m\omega^2}{2} \phi_j (2\phi_j - \phi_{j+1} - \phi_{j-1}) \right] \quad (9.8)$$

where $[\phi_i, \pi_j] = \delta_{ij}$. The generator of the discrete translational symmetry (from one lattice point to another) is the crystal momentum, which hence commutes with the Hamiltonian. Translational symmetry always implies that we should Fourier transform, however, in the case of a discrete translational symmetry the Fourier transform becomes a Fourier series:

$$\phi_j = \frac{1}{\sqrt{N}} \sum_q e^{iqR_j} \phi_q, \quad \pi_j = \frac{1}{\sqrt{N}} \sum_q e^{iqR_j} \pi_q \quad (9.9)$$

where $R_j = ja$ is the site of the j^{th} vertex. By applying these Fourier series to the Hamiltonian we obtain

$$H = \sum_q \left[\frac{1}{2m} \pi_q \pi_{-q} + \frac{m\omega_q^2}{2} \phi_q \phi_{-q} \right] \quad (9.10)$$

$$\omega_q \equiv 4\omega^2 \sin^2 \left(\frac{qa}{2} \right)$$

We can diagonalise this by means of a unitary transformation

$$\begin{pmatrix} \pi_q \\ \phi_q \end{pmatrix} = \begin{pmatrix} -i\sqrt{\frac{m\omega_q \hbar}{2}} & i\sqrt{\frac{m\omega_q \hbar}{2}} \\ \sqrt{\frac{\hbar}{2m\omega_q}} & \sqrt{\frac{\hbar}{2m\omega_q}} \end{pmatrix} \begin{pmatrix} a_q \\ a_q^\dagger \end{pmatrix} \quad (9.11)$$

from which we obtain

$$H = H_{\text{CM}} + \sum_{q \neq 0} \hbar \omega_q \left(a_q^\dagger a_q + \frac{1}{2} \right) \quad (9.12)$$

where we have condensed the collective motion of the centre-of-mass into

$$H_{\text{CM}} = \frac{1}{2m} \pi_0^2 \quad (9.13)$$

Hence we see that the interaction between neighbouring atoms in a chain results in bosonic quasiparticles which are referred to as *phonons*.

9.1.2 *Thermodynamic Limit*

Generally it is quite useful to be able to convert sums into integrals and in some cases integrals into sums. An example of a case in which this is possible is if we are summing over the different momenta of phonons in a lattice; this is a *sum* because the momenta are discrete, however, in the thermodynamic limit the difference between adjacent momenta becomes smaller and smaller

because $\Delta q = \frac{2\pi}{L}$ where L is the length of the chain (or volume in higher dimensions). The thermodynamic limit corresponds to letting $L \rightsquigarrow \infty$ which allows us to replace

$$\sum_q [\dots] \rightsquigarrow L \int_0^{\frac{2\pi}{a}} \frac{dq}{2\pi} [\dots] \quad (9.14)$$

In d -dimensions this would be

$$\sum_{\mathbf{q}} [\dots] \rightsquigarrow L^d \int_0^{\frac{2\pi}{a}} \frac{d^d \mathbf{q}}{(2\pi)^d} [\dots] \quad (9.15)$$

An application hereof would be to calculate the zero-point energy of phonons in a three-dimensional harmonic crystal. Similarly to what we found in one-dimension the energy is

$$H = \sum_{\mathbf{q}} \hbar \omega_{\mathbf{q}} \left(\hat{n}_{\mathbf{q}} + \frac{1}{2} \right) \quad (9.16)$$

so the zero-point energy is

$$E_0 = \sum_{\mathbf{q}} \frac{\hbar \omega_{\mathbf{q}}}{2} \xrightarrow{\text{Thermodynamic Limit}} L^3 \int \frac{d^3 \mathbf{q}}{(2\pi)^3} \frac{\hbar \omega_{\mathbf{q}}}{2} \quad (9.17)$$

Solving this expression after substituting in for $\omega_{\mathbf{q}}$ we get

$$E_0 \approx N \hbar \omega_0 \times 1.19 \quad (9.18)$$

so the zero-point energy per unit cell is a finite number, but a macroscopic number for the entire material.

9.1.3 The Continuum Limit

By letting the lattice spacing go towards zero $a \rightarrow 0$ we increase the size of the Brillouin zone to infinity and we would treat q as a continuous variable. However, this limit brings with it a divergence; the zero point energy scales with length in one-dimension

$$E_0 = \frac{L \hbar c}{2\pi \varepsilon^2} \quad (9.19)$$

9.2 CONSERVED PARTICLES

A general second quantised many-body Hamiltonian can be written as

$$H = \int d^3x \left[\psi^\dagger(x) \left(-\frac{\hbar^2}{2m} \nabla^2 + U(x) \right) \psi(x) + \frac{1}{2} \int d^3x' V(x-x') : \rho(x) \rho(x') : \right]$$

where we have introduced the *normal ordering operator*, which takes an expression and tells you that you need to convert it into a form where the annihilation operators are on the right, and the creation operators are on the left. For instance

$$:\psi(x)\psi(y)^\dagger: = \zeta \psi(y)^\dagger \psi(x) \quad (9.20)$$

where ζ is one for bosons and minus one for fermions.

9.2.1 *Coleman's Notation*

Henceforth Coleman uses a notation that needs some getting used to, but it simplifies expressions significantly and is quite well defined. Let $1 = (\mathbf{x}_1, \mathbf{k}_1, t_1, \sigma_1)$ i.e all relevant numbers that we would index with a 1. This allows us to write

$$[\psi(1), \psi(2)]_{\pm} = [\psi^{\dagger}(1), \psi^{\dagger}(2)]_{\pm} = 0, \quad [\psi(1), \psi^{\dagger}(2)]_{\pm} = \delta(1 - 2) \quad (9.21)$$

where $[\cdot, \cdot]_{\pm}$ is the anti-commutator for bosons (+) and the commutator for fermions (-).

9.2.2 *Field Operators in Different Bases*

Suppose we have two bases $\{|r\rangle\}$ and $\{|\tilde{s}\rangle\}$. The (unitary) transformation that takes us from one basis to the other:

$$|\tilde{s}\rangle = \sum_r \langle r|\tilde{s}\rangle |r\rangle \quad (9.22)$$

tells us about the transformation between field operators:

$$a_{\tilde{s}} = \langle \tilde{s}|\psi\rangle = \sum_r \langle \tilde{s}|r\rangle \langle r|\psi\rangle \quad (9.23)$$

This also works for continuous bases, such as the position and momentum bases:

$$\psi(x) = \int dq \langle x|q\rangle c_q = \frac{1}{\sqrt{L}} \int dq e^{-iqx} c_q \quad (9.24)$$

where $\langle x|q\rangle$ is normalised with $L^{-\frac{1}{2}}$ in one dimension. All observables are hermitian operators, which implies there exists a basis in which they are diagonal, thus for example Hamiltonian operators can always be diagonalised;

$$H = \sum_{\ell} \varepsilon_{\ell} \psi_{\ell}^{\dagger} \psi_{\ell} \quad (9.25)$$

we will quite often see that this diagonal basis is the momentum basis; in cases where the Hamiltonian is translationally invariant, however, this is not always the case and sometimes it is quite difficult to find the basis in which a Hamiltonian (or any operator) is diagonal.

Bogoliubov Transformation

Suppose we have **bosonic** operators a and a^{\dagger} , such that $[a, a^{\dagger}] = 1$, if we change basis

$$b = ua + va^{\dagger} \quad (9.26)$$

for $u, v \in \mathbb{C}$, it must be the case that

$$[b, b^{\dagger}] = (|u|^2 - |v|^2) [a, a^{\dagger}] = 1 \quad (9.27)$$

hence that $|u|^2 - |v|^2 = 1$, this is satisfied by

$$u = e^{i\theta_1} \cosh \beta, \quad v = e^{i\theta_2} \sinh \beta \quad (9.28)$$

where θ_1 and θ_2 are free variables.

Similarly for **fermionic** operators, where we have $\{a, a^\dagger\} = 1$, we require that

$$|u|^2 + |v|^2 = 1 \quad (9.29)$$

which is satisfied for

$$u = e^{i\theta_1} \cos \phi, \quad v = e^{i\theta_2} \sin \phi \quad (9.30)$$

again with θ_1 and θ_2 as free variables.

9.2.3 Field Operators as Creation and Annihilation Operators

We previously discussed the Fock Space, and that creation and annihilation operators can take us from one state in Fock Space to another, this is naturally true in all bases, which means we should denote the Fock State slightly differently from what we did previously. The only relevant quantity is the number of fermions or bosons in each state:

$$|n_1, n_2, \dots, n_\ell, \dots\rangle = \prod_\ell \frac{(\psi_\ell^\dagger)^{n_\ell}}{\sqrt{n_\ell!}} |0\rangle, \quad \text{bosons} \quad (9.31)$$

$$|n_1, n_2, \dots, n_r\rangle = (\psi_r^\dagger)^{n_r} \dots (\psi_1^\dagger)^{n_1} |0\rangle, \quad \text{fermions} \quad (9.32)$$

for bosons the numbers n_ℓ are non-negative integers, however, for fermions the Pauli exclusion principle prevents there from being more than one fermion in the same quantum state, which implies that the fermionic occupation numbers can only be $n_r \in \{0, 1\}$.

The wave function of the N -particle state $\Psi_S(t)$ is the overlap with:

$$\Psi_S(x_1, x_2, \dots, x_N, t) = \langle x_1, x_2, \dots, x_N | \Psi_S(t) \rangle = \langle 0 | \psi(x_1) \psi(x_2) \dots \psi(x_N) | \Psi_S(t) \rangle \quad (9.33)$$

note that the statistics is already encoded into this wavefunction due to the fact that we enforced the (anti-)commutation relations for the field operators. This means we never need to think about Slater-determinants again.

9.2.4 Interactions

Classically the interaction potential of a continuous plasma of particles is given by

$$V = \frac{1}{2} \int d^3x d^3x' V(x - x') \rho(x) \rho(x') \quad (9.34)$$

however if we were to just put hats on the densities, then one-particle states would interact with themselves. Thus, the correct form of the second quantised interaction is the above form minus the self interactions:

$$\hat{V} = \frac{1}{2} \int d^3x d^3x' V(x - x') \hat{\rho}(x) \hat{\rho}(x') - \frac{1}{2} \int d^3x V(0) \hat{\rho}(x) = \frac{1}{2} \int d^3x d^3x' V(x - x') : \hat{\rho}(x) \hat{\rho}(x') : \quad (9.35)$$

It is worth mentioning that

$$\hat{V} |x_1, x_2, \dots, x_N\rangle = \sum_{i < j} V(x_i - x_j) |x_1, x_2, \dots, x_N\rangle \quad (9.36)$$

so the state $|x_1, x_2, \dots, x_N\rangle$ is an eigenstate of the second quantised interaction potential, with eigenvalues of the classical interaction. Writing this in the momentum basis gives us quite a bit more insight. That is, we would like to Fourier transform

$$\hat{V} = \frac{1}{2} \int d^3\mathbf{x} d^3\mathbf{x}' V(\mathbf{x} - \mathbf{x}') \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}^\dagger(\mathbf{x}') \hat{\psi}(\mathbf{x}') \hat{\psi}(\mathbf{x}) \quad (9.37)$$

Using that

$$\psi_\sigma(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} c_{\mathbf{k}\sigma} e^{i\mathbf{k}\cdot\mathbf{x}} \quad (9.38)$$

And Fourier transforming the potential, $V(\mathbf{x} - \mathbf{x}') = \int \frac{d^3\mathbf{q}}{(2\pi)^3} V(\mathbf{q}) e^{i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')}$, we obtain

$$\hat{V} = \frac{1}{2} \int \frac{d^3\mathbf{k}_1}{(2\pi)^3} \frac{d^3\mathbf{k}_2}{(2\pi)^3} \frac{d^3\mathbf{q}}{(2\pi)^3} V(\mathbf{q}) c_{\mathbf{k}_1+\mathbf{q}}^\dagger c_{\mathbf{k}_2-\mathbf{q}}^\dagger c_{\mathbf{k}_2} c_{\mathbf{k}_1} \quad (9.39)$$

however, if we would include the possibility that the two interacting particles have different spin this would become

$$\hat{V} = \frac{1}{2} \sum_{\sigma\sigma'} \int_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}} V(\mathbf{q}) c_{\mathbf{k}_1+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}_2-\mathbf{q}, \sigma'}^\dagger c_{\mathbf{k}_2, \sigma'} c_{\mathbf{k}_1, \sigma} \quad (9.40)$$

Thus this is a process where we begin with two particles; with momentum \mathbf{k}_1 and \mathbf{k}_2 respectively, that are scattered by the potential to particles with momentum $\mathbf{k}_1 + \mathbf{q}$ and $\mathbf{k}_2 - \mathbf{q}$ respectively. We can represent this interaction with the following Feynman Diagram:

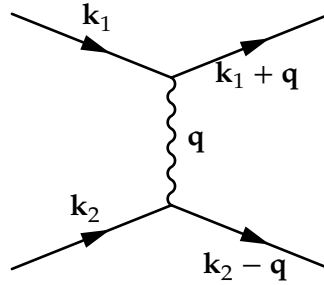


Figure 9.1: Feynman Diagram for a simple scattering process

9.3 SIMPLE EXAMPLE OF SECOND QUANTISATION

9.3.1 Jordan-Wigner Transformation

Let us now consider a spin- $\frac{1}{2}$ system, which will soon become a system containing many spin- $\frac{1}{2}$ particles. If we think of the spin-down state as $|0\rangle$ and the spin-up state as $|1\rangle$ then we can construct raising and lowering operators that take us from one to the other

$$|0\rangle = f|1\rangle, \quad |1\rangle = f^\dagger|0\rangle \quad (9.41)$$

Using this we can construct one-particle spin operators

$$S_x = \frac{1}{2}(f^\dagger + f), \quad S_y = \frac{1}{2i}(f^\dagger - f) \quad (9.42)$$

these are just the spin-raising and lowering operators we know from other Quantum Mechanics courses. However, an issue arises when we want to look at multi-particle states; in these the fermionic operators from different Hilbert spaces *anti*-commute, whereas the spin operators commute. The solution to this is described by the Jordan-Wigner transformation where we add an additional phase operator:

$$S_j^z = f_j^\dagger f_j - \frac{1}{2}, \quad S_j^+ = f_j^\dagger e^{i\pi\hat{\phi}_n}, \quad S_j^- = f_j e^{-i\pi\hat{\phi}_n} \quad (9.43)$$

where

$$\hat{\phi}_n \equiv \sum_{\ell < j} \hat{n}_\ell \quad (9.44)$$

Some useful identities here are that

$$\{f_j, e^{i\pi\hat{n}_j}\} = 0, \quad f_j e^{i\pi\hat{n}_j} = -f_j \quad (9.45)$$

$$\{e^{i\hat{\phi}_j}, f_\ell^{(\dagger)}\} = 0, \quad \text{for } l < j \quad (9.46)$$

$$[e^{i\hat{\phi}_j}, f_\ell^{(\dagger)}] = 0, \quad \text{for } l \geq j \quad (9.47)$$

where $f_\ell^{(\dagger)}$ means the identity works both for the raising and lowering operators. This transformation allows us to write magnetic Hamiltonians in forms that make them far easier to read. For instance the one-dimensional Heisenberg model,

$$H = -J \sum (S_j^x S_{j+1}^x + S_j^y S_{j+1}^y) - J_z \sum S_j^z S_{j+1}^z \quad (9.48)$$

can be written as

$$H = \sum_k \omega_k s_k^\dagger s_k - J_z \sum_j n_j n_{j+1} \quad (9.49)$$

where also have Fourier transformed and defined

$$\omega_k = J_z - J \cos ka \quad (9.50)$$

9.3.2 Hubbard Model

The nearest-neighbour Hubbard model describes localised electrons on a lattice, that can hop from one site to one of its neighbouring sites. By beginning with the real-space Hamiltonian that describes this, together with electron-electron interactions and Fourier transforming we obtained the Hubbard model Hamiltonian

$$H = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} c_{\mathbf{k}} \sigma^\dagger c_{\mathbf{k}} \sigma + \frac{U}{N} \sum_{\mathbf{q}, \mathbf{k}, \mathbf{k}'} c_{\mathbf{k}-\mathbf{q}} \uparrow^\dagger c_{\mathbf{k}'+\mathbf{q}, \downarrow}^\dagger c_{\mathbf{k}', \downarrow} c_{\mathbf{k}, \uparrow} \quad (9.51)$$

where

$$\varepsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y + \cos k_z) + \varepsilon \quad (9.52)$$

and t is the hopping constant and ε is the energy associated with being on one of the sites. Additionally U is the electron-electron interaction energy at the same site (hence the requirement for the spins to be opposite).

9.3.3 Fluid of non-interacting fermions, thermal equilibrium

The thermodynamics of a fluid of fermions is described by Fermi-liquid-theory, which for instance can be used to describe metals.

The free energy is a central quantity as we can use it to calculate different (measurable) quantities, such as entropy, pressure and density. For a single particle the free energy is described by the *single-free-energy-functional*

$$F = -2k_B T V \int_{\mathbf{k}} \ln [1 + \exp (-\beta(E_{\mathbf{k}} - \mu))] \quad (9.53)$$

In thermal equilibrium the number of fermions in a state with momentum $\mathbf{p} = \hbar \mathbf{k}$ is given by the Fermi-Dirac distribution

$$n_{\mathbf{k}} = f_{\text{FD}}(E_{\mathbf{k}} - \mu), \quad f_{\text{FD}}(x) = \frac{1}{e^{\beta x} + 1} \quad (9.54)$$

the low-temperature limit of the Fermi-Dirac distribution is a Heaviside $\Theta(\mu - E_{\mathbf{k}})$. In this limit the ground state is the state where all electrons are below the Fermi-level ($\varepsilon_F = \lim_{T \rightarrow 0} \mu(T) = \frac{\hbar k_F^2}{2m}$):

$$|\psi_G\rangle = \prod_{|\mathbf{k}| < k_F, \sigma} c_{\mathbf{k}\sigma}^\dagger |0\rangle \quad (9.55)$$

The ground state density of a Fermi gas is given by the number of occupied states divided by the volume of the Fermi-surface:

$$\langle \hat{\rho} \rangle = \frac{1}{V} \sum_{\mathbf{k}\sigma} \langle c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \rangle = 2 \int_{|\mathbf{k}| < k_F} \frac{d^3 \mathbf{k}}{(2\pi)^3} = \frac{2}{(2\pi)^3} V_{\text{FS}} \quad (9.56)$$

where

$$V_{\text{FS}} = \frac{4\pi k_F^3}{3} = \left(\frac{4\pi}{3} \right) \left(\frac{3m\varepsilon_F}{\hbar^2} \right)^{\frac{3}{2}} \quad (9.57)$$

hence

$$\langle \hat{\rho} \rangle = \frac{1}{3\pi^2} \left(\frac{3m\varepsilon_F}{\hbar^2} \right)^{\frac{3}{2}} \quad (9.58)$$

Some other important quantities are

$$P = -\frac{\partial F}{\partial V} = -2k_B T \int_{\mathbf{k}} \ln [1 + \exp (-\beta(E_{\mathbf{k}} - \mu))] \quad (9.59)$$

$$N = -\frac{\partial F}{\partial \mu} = 2 \int_{\mathbf{k}} f(E_{\mathbf{k}} - \mu) \quad (9.60)$$

9.4 GREEN'S FUNCTIONS

9.4.1 Interaction Picture

Picture	States	Operators
S	$i\partial_t \psi_S\rangle = H \psi_S\rangle$	Constant
H	Constant	$-i\partial_t O_H = [H, O_H]$
I	$i\partial_t \psi_I\rangle = V_I \psi_I\rangle$	$-i\partial_t O_I = [H_0, O_I]$

As shown in the table above, it is both operators and states that evolve in the interaction picture, however, the states evolve only according to the perturbation, V_I , whereas operators evolve according to the unperturbed Hamiltonian. Another way of seeing this is that we have moved all trivial time evolution on to the operators, such that we only need to worry about the potential's action on the state.

The time-evolution operator in the interaction picture is $U_I(t) = e^{iH_0 t} e^{-iHt}$, which has the following action

$$|\psi_I(t)\rangle = U_I(t)|\psi_I(0)\rangle \quad (9.61)$$

more generally

$$|\psi_I(t)\rangle = S_I(t, t')|\psi_I(t')\rangle, \quad S_I(t, t') = U_I(t)U_I^\dagger(t') \quad (9.62)$$

The time-evolution operator $S_I(t, t')$ will be used shortly to generate free-single-particle Green's functions, so this operator will play a central role in this section. Note that because $i\partial_t U_I(t) = V_I(t)U_I(t)$ we have that

$$i\partial_{t_2} S_I(t_2, t_1) = V_I(t_2)S_I(t_2, t_1) \quad (9.63)$$

9.4.2 Time-ordering operator

The time-ordering operator, T , allows us to write time-evolution operators in a very compact way. The time-ordering operator takes a product of time-dependent operators and orders them such that *later times are to the right*, such that when the operators operate on a ket, they will operate chronologically:

$$T\{O_1(t_1)O_2(t_2)\cdots O_N(t_N)\} = \zeta^p O_{p_1}(t_{p_1})\cdots O_{p_N}(t_{p_N}) \quad (9.64)$$

where again ζ is minus one for fermions and positive one for bosons and $t_{p_N} < t_{p_{N-1}} < \cdots < t_{p_1}$ and p is the parity of the permutation (whether it's an even or an odd permutation).

Using the time-ordering operator we can now express the time-evolution operator quite simply:

$$S_I(t_2, t_1) = T\left[\exp\left(-i\int_{t_1}^{t_2} dt V_I(t)\right)\right] \quad (9.65)$$

9.4.3 Driven harmonic oscillator

Consider now a (bosonic) harmonic oscillator

$$H_0 = \omega \left(b^\dagger b + \frac{1}{2} \right) \quad (9.66)$$

to which we apply the following perturbation

$$V(t) = \bar{z}(t)b + b^\dagger z(t) \quad (9.67)$$

where $z, \bar{z} \in \mathbb{C}$ are independent complex functions of time. In the case that we apply a force $H = H_0 - f(t)\hat{x}$ the z functions would be equal:

$$z(t) = \bar{z}(t) = \frac{f(t)}{\sqrt{2m\omega}} \quad (9.68)$$

The inner product between the ground state at $t = -\infty$ and $t = \infty$ defines a functional of z and \bar{z} , and is related to the probability of the perturbation exciting the ground state. This functional is given by

$$\Delta[\bar{z}, z] = \langle -\infty | \infty \rangle = \langle 0 | U_I(\infty) U_I^\dagger(-\infty) | 0 \rangle \quad (9.69)$$

which, using the time-ordered expression we obtained previously can be written as

$$\Delta[\bar{z}, z] = \left\langle T \exp \left(-i \int_{-\infty}^{\infty} dt [\bar{z}(t)b(t) + b^\dagger(t)z(t)] \right) \right\rangle_0 \quad (9.70)$$

A general property of time-evolution property is the propagation identity:

$$S(t_2, t_1) = S(t_2, t_i) S(t_i, t_1) \quad (9.71)$$

Thus we can break $S(\infty, -\infty)$ up into small bites propagating from some t to $t + \Delta t$:

$$\begin{aligned} S(\infty, -\infty) &= S(\infty, \infty - \Delta t) S(\infty - \Delta t, \infty - 2\Delta t) \\ &\dots S(\Delta t, 0) S(0, -\Delta t) \dots S(-\infty + \Delta t, -\infty) \end{aligned} \quad (9.72)$$

Over these small intervals of time we can approximate the integral $\int_t^{t+\Delta t} f(t') dt' \approx f(t)\Delta t$. Let us consider $S(t, -t)$ and split it up into N pieces as described above, then we let $N \rightarrow \infty$ and $t \rightarrow \infty$:

$$S(t, -t) = e^{A_N - A_N^\dagger} e^{A_{N-1} - A_{N-1}^\dagger} \dots e^{A_1 - A_1^\dagger} \quad (9.73)$$

where

$$A_j \equiv -i\bar{z}(t_j)b(t_j)\Delta t, \quad t_j \equiv t \left(\frac{2j}{N} - 1 \right) \quad (9.74)$$

$$A_j^\dagger \equiv iz(t_j)b^\dagger(t_j)\Delta t \quad (9.75)$$

due to the fact they A_j and A_j^\dagger are operators they will not commute with each other, so normal ordering Equation 9.73 is not quite trivial, however, we can use

$$e^{\hat{a} + \hat{\beta}} = e^{\hat{\beta}} e^{\hat{a}} e^{[\hat{a}, \hat{\beta}]/2}, \quad e^{\hat{a}} e^{\hat{\beta}} = e^{\hat{\beta}} e^{\hat{a}} e^{[\hat{a}, \hat{\beta}]} \quad (9.76)$$

which only holds if $[\hat{\alpha}, [\hat{\alpha}, \hat{\beta}]] = [\hat{\beta}, [\hat{\alpha}, \hat{\beta}]] = 0$. Using this we can write

$$S(t, -t) = e^{-\sum_j A_j^\dagger} e^{\sum_j A_j} \exp\left(-\sum_{j \geq k} [A_j, A_k^\dagger] \left(1 - \frac{1}{2} \delta_{jk}\right)\right) \quad (9.77)$$

The commutator

$$[A_j, A_k^\dagger] = \bar{z}(t_j) z(t_k) \Delta t^2 e^{-i\omega(t_j - t_k)} \quad (9.78)$$

where we've used that H_0 is the simple harmonic oscillator which means that $b(t) = be^{-i\omega t}$. Note that $[A_j, A_k^\dagger] \in \mathbb{C}$, so the operator part of $S(t, -t)$ is in the first two exponentials. However, note that the lowering operator on the vacuum state gives zero, so

$$e^{\sum_j A_j} |0\rangle = e^{\sum_j 0} |0\rangle = |0\rangle \quad (9.79)$$

we are taking the expectation value with respect to the vacuum state $\langle \cdot \rangle_0$.

therefore $\langle e^{-\sum_j A_j^\dagger} e^{\sum_j A_j} \rangle_0 = 1$:

$$\begin{aligned} \langle S(t, -t) \rangle_0 &= \\ \exp\left(-\sum_{j,k} \Delta t^2 \bar{z}(t_j) \Theta(t_j - t_k) e^{-i\omega(t_j - t_k)} z(t_k) (1 - \delta_{jk})\right) \end{aligned} \quad (9.80)$$

in the limit where $N \rightarrow \infty$ and hence $\Delta t \rightarrow 0$ we obtain integrals, however, the term with δ_{jk} only has one sum and a Δt^2 , which implies that it dies in the limit where $\Delta t \rightarrow 0$. Thus we are left with

$$\begin{aligned} \langle S(t, -t) \rangle_0 &= \\ \exp\left(-\int_{-t}^t dt' dt'' \bar{z}(t') \Theta(t' - t'') e^{-i\omega(t' - t'')} z(t'')\right) \end{aligned} \quad (9.81)$$

we are interested in the limit where $t \rightarrow \infty$, in which case we obtain our generating functional:

$$\Delta[\bar{z}, z] = \exp\left(-\int_{-\infty}^{\infty} dt' dt'' \bar{z}(t') G(t' - t'') z(t'')\right) \quad (9.82)$$

where $G(t' - t'')$ is our first example of a Green's function:

$$G(t' - t'') \equiv \Theta(t' - t'') e^{-i\omega(t' - t'')} \quad (9.83)$$

9.4.4 Wick's theorem

We now have two identities that describe $\Delta[\bar{z}, z]$: Equations 9.70 and 9.82. If we expand up to linear order in \bar{z} and z , we obtain:

$$\begin{aligned} (-i)^2 \int_{-\infty}^{\infty} dt' dt'' \bar{z}(t') \langle T \{b(t') b^\dagger(t'')\} \rangle_0 z(t'') + O(\bar{z}^2, z^2) \\ = -i \int_{-\infty}^{\infty} dt' dt'' \bar{z}(t') G(t' - t'') z(t'') \end{aligned} \quad (9.84)$$

hence

$$G(t' - t'') = -i \langle T \{b(t') b^\dagger(t'')\} \rangle_0 \quad (9.85)$$

note that the terms $\langle T \{b(t)\} \rangle_0$ and $\langle T \{b^\dagger(t)\} \rangle_0$ are both zero, because of the fact that we are taking an expectation value. Similarly, to second order we would have

$$G(t_1, t_2, t_3, t_4) = \langle T \{b(t_1) b(t_2) b^\dagger(t_3) b^\dagger(t_4)\} \rangle_0 \quad (9.86)$$

and so on. Wick's theorem tells us how we can take these many-body Green's functions and decompose them into products of single-particle Green's functions. **Wick's Theorem:**

$$G(1, \dots, n; 1', \dots, n') = \sum_P \prod_r \zeta^P G(r - P'_r) \quad (9.87)$$

where the ζ has been included so that this result also holds for fermions. For fermions the derivation is very similar, except that we must treat the source and sink terms $\bar{z} \rightarrow \bar{\eta}$ and $z \rightarrow \eta$ as *Grassman numbers*, which anti-commute with each other and with field operators.

In reality Wick's theorem is more general than what is given above; Wick's theorem refers to general products of operators that can be contracted by means of the time-ordering and normal-ordering operators. However, in this course we only used Wick's theorem in the context of Green's functions (and later response functions).

9.4.5 Gell-Mann-Low Theorem

The Gell-Mann-Low theorem relates a time-ordered expectation value in the Heisenberg picture to a time-ordered expectation value in the interaction picture, provided that the perturbation is turned on and off adiabatically:

$$\langle \phi | T \{ A(t_1) B(t_2) \dots \} | \phi \rangle_H = \frac{\langle -\infty | T S(\infty, -\infty) A(t_1) B(t_2) \dots | -\infty \rangle_I}{\langle -\infty | S(\infty, -\infty) | -\infty \rangle} \quad (9.88)$$

where the state $| -\infty \rangle$ is well known $| -\infty \rangle_H = | -\infty \rangle_I$. Using this we can relate full interacting Green's functions to non-interaction Green's functions.

9.4.6 Useful quantities calculable from Green's functions

The expectation value of the density operator can be written as a Green's function:

$$\langle \phi | \hat{\rho}(\mathbf{x}, t) | \phi \rangle = \sum_{\sigma} \langle \phi | \hat{\psi}_{\sigma}^{\dagger}(\mathbf{x}, t) \hat{\psi}_{\sigma}(\mathbf{x}, t) | \phi \rangle = -(2S + 1) \lim_{\delta \rightarrow 0^+} G(\mathbf{x} = \mathbf{0}, t = -\delta) = -i(2S + 1) G(\mathbf{0}, 0^-) \quad (9.89)$$

Similarly, we can calculate the expectation value of the kinetic energy:

$$\begin{aligned} \langle \phi | \hat{T}(\mathbf{x}) | \phi \rangle &= - \sum_{\sigma} \frac{\hbar^2}{2m} \langle \phi | \psi_{\sigma}^{\dagger}(\mathbf{x}) \nabla_{\mathbf{x}}^2 \psi_{\sigma}(\mathbf{x}) | \phi \rangle \\ &= \frac{\hbar^2 \nabla_{\mathbf{x}}^2}{2m} (2S + 1) \langle \phi | T \psi(\mathbf{x}, 0^-) \psi^{\dagger}(\mathbf{x}', 0) | \phi \rangle \Big|_{\mathbf{x}=\mathbf{x}'} \\ &= i(2S + 1) \frac{\hbar^2 \nabla^2}{2m} G(\mathbf{x}, 0^-) \Big|_{\mathbf{x}=0} \end{aligned} \quad (9.90)$$

9.5 ZERO-TEMPERATURE FEYNMAN DIAGRAMS

The generating function for fermions $\mathcal{Z}[\bar{\eta}, \eta]$ can be written diagrammatically as

$$\mathcal{Z}[\bar{\eta}, \eta] = \exp \left(-i \bar{\eta} \bullet \longleftarrow \bullet \eta \right) \quad (9.91)$$

where we have associated

$$\int d1d2 \bar{\eta}(2)G(2-1)\eta(1) = \bar{\eta} \bullet \longleftarrow \bullet \eta \quad (9.92)$$

Note that

$$i \frac{\delta^2 \ln \mathcal{S}[\bar{\eta}, \eta]}{\delta \eta(2) \delta \bar{\eta}(1)} = G(1-2) = 1 \bullet \longleftarrow \bullet 2 \quad (9.93)$$

A general expectation value of the form

$$\langle \phi | T \mathcal{S}[\bar{\eta}, \eta] F[\psi^\dagger, \psi] | \phi \rangle \quad (9.94)$$

can be shown to be equal to

$$F \left[i \zeta \frac{\delta}{\delta \eta}, i \frac{\delta}{\delta \bar{\eta}} \right] \exp(-i \bar{\eta} \bullet \longleftarrow \bullet \eta) \quad (9.95)$$

if we now set $F[\psi^\dagger, \psi] = T \exp(-i \int dt V[\psi^\dagger, \psi])$ we obtain the *interacting generator*:

$$\mathcal{S}_I[\bar{\alpha}, \alpha] = \exp \left(i^{n-1} \int_{-\infty}^{\infty} dt V \left[\zeta \frac{\delta}{\delta \alpha}, \frac{\delta}{\delta \bar{\alpha}} \right] \right) \exp(-i \bar{\eta} \bullet \longleftarrow \bullet \eta) \quad (9.96)$$

where $\alpha = \eta$ and $\bar{\alpha} = -i\bar{\eta}$.

In processes where two particles scatter of each other at once, such as the Coulomb interaction, we have that

$$i^{n-1} V \left[\zeta \frac{\delta}{\delta \alpha}, \frac{\delta}{\delta \bar{\alpha}} \right] = \frac{i}{2} \int d^3x d^3x' V(x-x') \frac{\delta}{\delta \alpha(x)} \frac{\delta}{\delta \alpha(x')} \frac{\delta}{\delta \bar{\alpha}(x')} \frac{\delta}{\delta \bar{\alpha}(x)} \quad (9.97)$$

where $iV(x-x')$ is the two-particle scattering amplitude. The scattering process is depicted by

$$\begin{array}{ccc} \frac{\delta}{\delta \alpha(1)} & & \frac{\delta}{\delta \alpha(2)} \\ & \nwarrow \quad \nearrow & \\ \frac{1}{2} & \text{---} & \\ & \swarrow \quad \searrow & \\ \frac{\delta}{\delta \bar{\alpha}(1)} & & \frac{\delta}{\delta \bar{\alpha}(2)} \end{array}$$

Which gives us

$$\mathcal{S}_I[\bar{\alpha}, \alpha] = \exp \left(\begin{array}{ccc} \frac{\delta}{\delta \alpha(1)} & & \frac{\delta}{\delta \alpha(2)} \\ & \nwarrow \quad \nearrow & \\ & \text{---} & \\ & \swarrow \quad \searrow & \\ \frac{\delta}{\delta \bar{\alpha}(1)} & & \frac{\delta}{\delta \bar{\alpha}(2)} \end{array} \right) \exp(-i \bar{\eta} \bullet \longleftarrow \bullet \eta)$$

This should be thought of in terms of the exponential function's power series. The lowest nonzero (non-trivial) term is

$$\frac{1}{2} \left(\begin{array}{ccc} \frac{\delta}{\delta \alpha(1)} & & \frac{\delta}{\delta \alpha(2)} \\ & \nwarrow \quad \nearrow & \\ & \text{---} & \\ & \swarrow \quad \searrow & \\ \frac{\delta}{\delta \bar{\alpha}(1)} & & \frac{\delta}{\delta \bar{\alpha}(2)} \end{array} \right) (-i \bar{\eta} \bullet \longleftarrow \bullet \eta)^2 \quad (9.98)$$

$$= \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} \quad (9.99)$$

9.5.1 *Linked-Cluster Theorem*

The linked-cluster theorem states that

$$\ln \mathcal{Z}_I[\bar{\alpha}, \alpha] = \sum \{\text{linked - cluster diagrams}\} \quad (9.100)$$

which can be shown by making identical(commuting) copies of \mathcal{Z}_I and using the identity that

$$\ln \mathcal{Z}_I = \lim_{n \rightarrow 0} \left[\frac{\mathcal{Z}_I^n - 1}{n} \right] \quad (9.101)$$

The linked cluster diagrams are all connected diagrams, so the dumbbell and oyster diagram are of order α^0 , next are single particle propagators with α and $\bar{\alpha}$ at the ends, then come two-particle propagators etc. But there are *not* products of diagrams, as these are not linked.

A similar statement can be made about the full two-particle Green's function:

$$G(2-1) = \frac{\delta^2 \ln \mathcal{Z}_I[\bar{\alpha}, \alpha]}{\delta \alpha(1) \delta \bar{\alpha}(2)} \bigg|_{\alpha, \bar{\alpha}=0} = \sum \{\text{two - leg diagrams}\} \quad (9.102)$$

and similarly

$$G(1,2;3,4) = \sum \{\text{four - legged diagrams}\} \quad (9.103)$$

9.5.2 *Feynman Rules in Momentum Space*



1. Fermions with four-momentum, k , and spin σ :  $\equiv G_\sigma^0(k)$
2. Interaction lines with four-momentum, q :  $\equiv iV(q)$
3. Conserve spin and four-momentum at each vertex; incoming and outgoing momenta are equal and spins cannot be flipped
4. Multiply by $(-1)^F (2S+1)^F$ where F is the number of fermion loops.
5. If there are any "same-time" fermion-loops, i.e ones that loop back on themselves, include (a) convergence factor(s), $e^{ik_n 0^+}$
6. For each *internal* four-momentum, k , integrate $\int \frac{d^4 k}{(2\pi)^4}$
7. Divide by the symmetry factor, p , of the diagram and multiply by VT (volume times time-interval)

Table 9.1: *Feynman Rules in Momentum Space, as per Bruus & Flensberg*

9.5.3 *Hartree-Fock energy*

The Hartree-Fock shift in the ground-state energy of an interacting electron gas to first order is given by

$$E_g = E_0 + i\mathcal{U} \left(\text{diagram 1} + \text{diagram 2} \right) \quad (9.104)$$

Using Feynman's momentum-space rules we can write the energy shift out, noting that $p = 2$ for both diagrams:

$$\Delta E_{\text{HF}} = \frac{i\mathcal{U}}{2} \int \frac{d^4k d^4k'}{(2\pi)^8} e^{i(\omega+\omega')\delta} \left[(-[2S+1])^2 (iV_{\mathbf{q}=\mathbf{0}}) + (-[2S+1]) (iV_{\mathbf{k}=\mathbf{k}'}) \right] G(k)G(k') \quad (9.105)$$

where $k = (\mathbf{k}, \omega)$ and similarly for k' . It is only the Green's functions that depend on the frequencies, so let us integrate those away

$$\langle c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \rangle = -i \int \frac{d\omega}{2\pi} G(k) e^{i\omega\delta} = f_{\mathbf{k}} = \Theta(k_F - |\mathbf{k}|) \quad (9.106)$$

giving us

$$\Delta E_{\text{HF}} = \frac{i\mathcal{U}}{2} (2S+1) \int_{\mathbf{k}\mathbf{k}'} \left[(2S+1) V_{\mathbf{q}=\mathbf{0}} - V_{\mathbf{k}-\mathbf{k}'} \right] f_{\mathbf{k}} f_{\mathbf{k}'} \quad (9.107)$$

Unfortunately the $\mathbf{q} = \mathbf{0}$ term is infinite, but the book argues that we can neglect this, which means the first order approximation of the ground-state energy of an electron gas with interactions taken into account can be written as

$$\frac{E_g}{\mathcal{U}} = (2S+1) \left(\int_{\mathbf{k}} \frac{\hbar^2 k^2}{2m} f_{\mathbf{k}} - \frac{1}{2} \int_{\mathbf{k}\mathbf{k}'} f_{\mathbf{k}} f_{\mathbf{k}'} \frac{e^2}{\epsilon_0 (\mathbf{k} - \mathbf{k}')^2} \right) \quad (9.108)$$

carefully evaluating these integrals gives us

$$\frac{E_g}{\rho\mathcal{U}} = \frac{3}{5} \epsilon_F - \frac{3}{4\pi} \frac{e^2 k_F}{4\pi\epsilon_0} \quad (9.109)$$

9.5.4 Exchange correlation

Due to the Pauli Exclusion Principle we expect that if we know an electron has a charge distribution $\rho(x)$, then another electron will have a low probability of being there where $\rho(x)$ is at its maximum. It is useful to consider the equal-time density correlation function:

$$C_{\sigma\sigma'}(\mathbf{x} - \mathbf{x}') = \langle \phi_0 | : \hat{\rho}_\sigma(\mathbf{x}) \hat{\rho}_{\sigma'}(\mathbf{x}') : | \phi_0 \rangle \quad (9.110)$$

It turns out that we can relate this to the Hartree-Fock energy (first order approximation of the correction due to interactions of the energy of an electron gas), because

$$\langle \phi_0 | V | \phi_0 \rangle = \frac{1}{2} \sum_{\sigma\sigma'} \int d^3\mathbf{x} d^3\mathbf{x}' V(\mathbf{x} - \mathbf{x}') \times \quad (9.111)$$

$$\langle \phi_0 | : \hat{\rho}_\sigma(\mathbf{x}) \hat{\rho}_{\sigma'}(\mathbf{x}') : | \phi_0 \rangle \\ = \frac{1}{2} \sum_{\sigma\sigma'} \int d^3\mathbf{x} d^3\mathbf{x}' V(\mathbf{x} - \mathbf{x}') C_{\sigma\sigma'}(\mathbf{x} - \mathbf{x}') \quad (9.112)$$

By removing the potential part of the dumbbell and oyster diagrams, we can hence obtain a diagrammatic expression for the correlation function

$$C_{\sigma\sigma'}(\mathbf{x} - \mathbf{x}') = - \left[\left(\begin{array}{c} \mathbf{x}, \sigma \quad \mathbf{x}', \sigma' \\ \text{[Diagram: Two separate circles, one at } \mathbf{x}, \sigma \text{ and one at } \mathbf{x}', \sigma' \text{]} \end{array} \right) + \left(\begin{array}{c} \mathbf{x} \quad \mathbf{x}', \sigma' \\ \text{[Diagram: A circle with a horizontal line connecting } \mathbf{x} \text{ and } \mathbf{x}', \sigma' \text{]} \end{array} \right) \delta_{\sigma\sigma'} \right] \quad (9.113)$$

which we can write out explicitly

$$C_{\sigma\sigma'}(\mathbf{x} - \mathbf{x}') = - \left[(-G(\mathbf{0}, 0^-))^2 - \right. \quad (9.114)$$

$$\left. \delta_{\sigma\sigma'} G(\mathbf{x} - \mathbf{x}', 0^-) G(\mathbf{x}' - \mathbf{x}, 0^-) \right] \\ = \rho_0^2 + \delta_{\sigma\sigma'} G(\mathbf{x} - \mathbf{x}', 0^-) G(\mathbf{x}' - \mathbf{x}, 0^-) \quad (9.115)$$

We can evaluate the Green's function

$$G(\mathbf{r}, 0^-) = \int_{\mathbf{k}} G(\mathbf{k}, 0^-) = i \int_{\mathbf{k}} f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (9.116)$$

$$= i \int \frac{dk}{2\pi^2} \frac{k \sin kr}{r} = i\rho_0 P(k_F r) \quad (9.117)$$

where

$$P(x) = 3 \left(\frac{\sin x - x \cos x}{x^3} \right) \quad (9.118)$$

which is an even function, therefore the two Green's functions are equal.

$$C_{\sigma\sigma'}(r) = \rho_0^2 \left(1 - \delta_{\sigma\sigma'} [P(k_F r)]^2 \right) \quad (9.119)$$

9.5.5 Self-energy

The full Green's function of a fermion in an interacting environment can be written as an infinite sum of terms that depend on the free Green's function to higher and higher order. However, it is possible to express this infinite sum in terms of a recurring term, the self-energy, which allows us to evaluate the sum exactly, just as one can do for geometric series. The self-energy is the sum of all scattering processes that cannot be separated into disconnected diagrams by cutting a single propagator.

Denote the full Green's function as

$$G(k) = \text{---} \Rightarrow \text{---} \quad (9.120)$$

and the self energy as

$$\Sigma(\mathbf{k}, \omega) = \text{---} \bigcirc \Sigma \text{---} \quad (9.121)$$

then

$$\text{---} \Rightarrow \text{---} = \text{---} \Rightarrow \text{---} + \text{---} \Rightarrow \bigcirc \Sigma \Rightarrow \text{---} \quad (9.122)$$

Which algebraically would be written as

$$G(\mathbf{k}, \omega) = G^0(\mathbf{k}, \omega) + G^0(\mathbf{k}, \omega) \Sigma(\mathbf{k}, \omega) G(\mathbf{k}, \omega) \quad (9.123)$$

which is a *Dyson Equation* and solving it yields

$$G(\mathbf{k}, \omega) = \frac{1}{\omega - \varepsilon_{\mathbf{k}} - \Sigma(\mathbf{k}, \omega)} \quad (9.124)$$

Generally, the self-energy can have be $\Sigma \in \mathbb{C}$, where the real part corresponds to a shift in the energy levels ("renormalised energy"), whereas the imaginary part gives the electrons a lifetime.

where the susceptibility tensor is given by

$$\left(\underline{\chi}\right)_{ij}(x) = i\mu_B^2 \langle [\sigma_i(x), \sigma_j(0)] \rangle_0 \quad (9.134)$$

which we can Fourier transform:

$$\chi_{ij}(q) = i\mu_B^2 \int d^4x \langle [\sigma_i(x), \sigma_j(0)] \rangle_0 \Theta(t) e^{-iq \cdot x} \quad (9.135)$$

The susceptibility is related to the time-ordered susceptibility through $\chi_{ij}(\mathbf{q}, \nu) = -i\chi_{ij}^T(\mathbf{q}, \nu + i\delta)$ where the latter is given by

$$\chi_{ij}^T(q) = \mu_B^2 \left[\sigma_j \begin{array}{c} \text{---} k \text{---} \\ \text{---} k+q \text{---} \end{array} \sigma_i \right] = -2\mu_B^2 \delta_{ij} \int_{\mathbf{k}} \left[\frac{1}{\omega + \nu - \tilde{\epsilon}_{\mathbf{k}+\mathbf{q}}} \frac{1}{\omega - \tilde{\epsilon}_{\mathbf{k}}} \right] \quad (9.136)$$

where $\tilde{\epsilon}_{\mathbf{k}} = \epsilon_{\mathbf{k}} - i\delta \text{sgn}(\epsilon_{\mathbf{k}})$ and $k = (\mathbf{k}, \omega)$. Performing the calculation gives us the *dynamical spin susceptibility*:

$$\chi(\mathbf{q}, \nu + i\delta) = 2\mu_B^2 \int_{\mathbf{k}} \frac{f_{\mathbf{k}+\mathbf{q}} - f_{\mathbf{k}}}{\nu - (\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}) + i\delta} \quad (9.137)$$

where $f_{\mathbf{k}}$ still is the zero-temperature limit of the Fermi-Dirac Distribution. In the $\nu \rightarrow 0$ limit we obtain the *static susceptibility*, which tells us about the response to a spatially (but not temporally) varying magnetic field, in which case we obtain

$$\chi(\mathbf{q}) = 2\mu_B^2 N(0) F\left(\frac{q}{2k_F}\right), \quad F(x) = \frac{1-x^2}{4x} \ln \left| \frac{1+x}{1-x} \right| + \frac{1}{2} \quad (9.138)$$

9.6 NON-ZERO-TEMPERATURE MANY-BODY PHYSICS

Non-zero-temperature many-body physics begins with the recognition that the time-evolution operator $U(t) = \exp(-itH)$ and the quantum mechanical partition function $\hat{\rho} \sim e^{-\beta H}$ are identical if we treat t as an imaginary number:

$$t \rightarrow -i\beta \quad (9.139)$$

Using this we can express the partition function in terms of the imaginary-time-evolution operator:

$$Z = \text{Tr } U(-i\beta) \quad (9.140)$$

and we can do the same with thermal-equilibrium expectation values:

$$\langle A \rangle = \frac{\text{Tr } [U(-i\beta)A]}{\text{Tr } [U(-i\beta)]} \quad (9.141)$$

It turns out that the machinery that we developed for zero-temperature many-body physics can be generalised to non-zero-temperature many-body physics, simply by letting $t \rightarrow -i\tau$ where the imaginary-time parametre $\tau \in [0, \beta]$. For instance, operators in the Heisenberg picture would be:

$$A_H(\tau) = e^{\tau H} A e^{-\tau H} \quad (9.142)$$

and the creation and annihilation operators in the free-particle Hamiltonian $H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}}$ are described by means of their Heisenberg equation of motion:

$$\frac{\partial c_{\mathbf{k}}}{\partial \tau} = [H, c_{\mathbf{k}}] = -\epsilon_{\mathbf{k}} c_{\mathbf{k}}, \quad \frac{\partial c_{\mathbf{k}}^\dagger}{\partial \tau} = [H, c_{\mathbf{k}}^\dagger] = \epsilon_{\mathbf{k}} c_{\mathbf{k}}^\dagger \quad (9.143)$$

which tells us that $c_k(\tau) = e^{-\varepsilon_k \tau} c_k$ and $c_k^\dagger(\tau) = e^{\varepsilon_k \tau} c_k^\dagger$. Note that even though $(c_k^\dagger)^\dagger = c_k$ it is *not* the case that $(c_k^\dagger(\tau))^\dagger = c_k(\tau)$ – i.e they are not hermitian conjugates.

We also have the (imaginary-)time-ordering operator, which has exactly the same effect as the regular time-ordering operator, which allows us to write

$$U(\tau) = T \exp \left[- \int_0^\tau d\tau V_I(\tau) \right] \quad (9.144)$$

$$S(\tau_2, \tau_1) = T \exp \left[- \int_{\tau_1}^{\tau_2} d\tau V_I(\tau) \right] \quad (9.145)$$

$S(\beta, 0)$ is of particular interest, because this tells us about the change in the Helmholtz free energy, and can be expressed in terms of the full partition function, Z , and the free partition function, Z_0 :

$$\frac{Z}{Z_0} = e^{-\beta \Delta F} = \left\langle T \exp \left[- \int_0^\beta d\tau V_I(\tau) \right] \right\rangle \quad (9.146)$$

In imaginary-time many-body physics it is the partition function that is the generator of (imaginary-time) Green's functions, which makes sense because we just saw that Z is related to $S(\beta, 0)$ (time-evolution operator over the entire τ interval), which corresponds to $S(\infty, -\infty)$ which was the generator for real-time Green's functions. We will denote imaginary time Green's functions as $\mathcal{G}(\tau)$.

9.6.1 Imaginary-Time Green's Functions

The imaginary-time Green's function is defined as

$$\mathcal{G}_{\lambda\lambda'}(\tau - \tau') = - \langle T \psi_\lambda(\tau) \psi_{\lambda'}^\dagger(\tau') \rangle = -\text{tr} \left[e^{-\beta(H-F)} \psi_\lambda(\tau) \psi_{\lambda'}^\dagger(\tau') \right] \quad (9.147)$$

where we have assumed that H is invariant under temporal translations, such that $\mathcal{G}(\tau, \tau') = \mathcal{G}(\tau - \tau')$, which is often the case. Additionally, it is often the case that the quantum number λ is conserved (for example very often spin is not flipped by H), in which case $\mathcal{G}_{\lambda\lambda'}(\tau - \tau') = \delta_{\lambda\lambda'} \mathcal{G}_\lambda(\tau - \tau')$.

For non-interacting particles $H = \sum_\lambda \xi_\lambda \psi_\lambda^\dagger \psi_\lambda$, where $\xi_\lambda = \varepsilon_\lambda - \mu$ the equal-time expectation value of the fields is

$$\langle \psi_\lambda^\dagger \psi_{\lambda'} \rangle = \delta_{\lambda\lambda'} \begin{cases} n(\xi_\lambda) & \text{Bosons} \\ f_{\text{FD}}(\xi_\lambda) & \text{Fermions} \end{cases} \quad (9.148)$$

and we have that $\langle \psi_{\lambda'} \psi_\lambda^\dagger \rangle = \delta_{\lambda\lambda'} - \langle \psi_\lambda^\dagger \psi_{\lambda'} \rangle$, which tells us that

$$\mathcal{G}_\lambda(\tau) = -e^{-\xi_\lambda \tau} \begin{cases} (1 + n(\xi_\lambda)) \Theta(\tau) + n(\xi_\lambda) \Theta(-\tau) & \text{B} \\ (1 - f_{\text{FD}}(\xi_\lambda)) \Theta(\tau) - f_{\text{FD}}(\xi_\lambda) \Theta(-\tau) & \text{F} \end{cases} \quad (9.149)$$

Due to the fact that our imaginary-time parametre, τ , spans from 0 to β the behaviour of imaginary-time Green's functions are slightly different from real-time Green's functions. For instance it can be shown that for $-\beta < \tau < 0$ we have

$$\mathcal{G}_\lambda(\tau + \beta) = \zeta \mathcal{G}_\lambda(\tau) \quad (9.150)$$

where as usual ζ is negative one for fermions and positive one for bosons

9.6.2 Matsubara Representation

The (anti-)periodicity of the Green's functions implies that we can decompose them into an infinite sum of complex exponentials. The *Matsubara Frequencies* are defined as

$$\nu_n = \frac{2\pi n}{\beta}, \quad \text{Bosons} \quad (9.151)$$

$$\omega_n = \frac{\pi(2n+1)}{\beta}, \quad \text{Fermions} \quad (9.152)$$

we choose these so that

$$e^{i\nu_n(\tau+\beta)} = e^{i\nu_n\tau} \quad (9.153)$$

$$e^{i\omega_n(\tau+\beta)} = -e^{i\omega_n\tau} \quad (9.154)$$

Now, using these we can decompose the Green's functions:

$$G_{\lambda\lambda'} = \frac{1}{\beta} \begin{cases} \sum_n G_{\lambda\lambda'}(i\nu_n) e^{-i\nu_n\tau} & \text{Bosons} \\ \sum_n G_{\lambda\lambda'}(i\omega_n) e^{-i\omega_n\tau} & \text{Fermions} \end{cases} \quad (9.155)$$

where

$$G_{\lambda\lambda'}(i\alpha_n) = \int_0^\beta d\tau G_{\lambda\lambda'}(\tau) e^{i\alpha_n\tau} \quad (9.156)$$

for some Matsubara frequency α_n . Using this we can calculate the Matsubara representation Green's functions for free particles:

$$G_\lambda(i\omega_n) = -\int_0^\beta d\tau e^{(i\omega_n - \xi_\lambda)\tau} (1 - f_{\text{FD}}(\xi_\lambda)) = \frac{-1}{i\omega_n - \xi_n} \frac{e^{(i\omega_n - \xi_\lambda)\beta} - 1}{e^{-\beta\xi_\lambda} + 1} = \frac{1}{i\omega_n - \xi_\lambda} \quad (9.157)$$

where we have used that $e^{i\omega_n\beta} = -1$. Similarly for bosons we get

$$G_\lambda(i\nu_n) = \frac{1}{i\nu_n - \xi_\lambda} \quad (9.158)$$

note how closely the zero and non-zero temperature Green's functions are related to each other.

9.6.3 Contour-Integral Method

The first important point to make is that the distribution functions have first order poles with weight $\frac{\xi}{\beta}$ at the Matsubara frequencies:

$$\lim_{z \rightarrow i\omega_n} (z - i\omega_n) f(z) = \lim_{z \rightarrow i\omega_n} \frac{z - i\omega_n}{1 + e^{z\beta}} \approx \lim_{z \rightarrow i\omega_n} \frac{z - i\omega_n}{\beta(i\omega_n - z)} = -\frac{1}{\beta} \quad (9.159)$$

and similarly for bosons. Now, consider a sum of the form

$$\frac{1}{\beta} \sum_n F(i\omega_n) \quad (9.160)$$

for some function $F(z)$ that tends to zero faster than $\frac{1}{|z|}$. Let γ be a circular path in the Argand plane with centre at the origin and radius ∞ . If the integrand converges towards zero quickly enough, then the contour integral along γ must be zero, which implies that the sum of all residues is zero. Suppose $F(z)$ has poles at z_i and $f_{\text{FD}}(z)$ has poles at $i\omega_n$ as previously established, then it must be the case that

$$0 = \oint_{\gamma} \frac{dz}{2\pi i} F(z) f_{\text{FD}}(z) = \sum_i \text{Res} [F(z_i) f_{\text{FD}}(z_i)] + \sum_n \text{Res} [F(i\omega_n) f_{\text{FD}}(i\omega_n)] \quad (9.161)$$

which gives us

$$\frac{1}{\beta} \sum_n F(i\omega_n) = \sum_i f_{\text{FD}}(z_i) \text{Res} [F(z_i)] \quad (9.162)$$

it is essential that $F(z)$ and $f_{\text{FD}}(z)$ don't have any common poles. However, quite often $F(z)$ will have poles along the real axis, whereas $f_{\text{FD}}(z)$ has poles on the imaginary axis, which avoids this problem.

Consider the following response function

$$\chi_0(\mathbf{q}, \tau) = -\frac{1}{\mathcal{V}} \langle T \rho(\mathbf{q}, \tau) \rho(-\mathbf{q}, 0) \rangle = \frac{1}{\mathcal{V}} \sum_{\mathbf{k}\sigma} \mathcal{G}_{0,\sigma}(\mathbf{k} + \mathbf{q}, \tau) \mathcal{G}_{0,\sigma}(\mathbf{k}, -\tau) \quad (9.163)$$

where we obtained Equation 9.163 with Wick's theorem. Now it is beneficial to Fourier transform, such that we are left with

$$\chi_0(\mathbf{q}, \nu_\ell) = \frac{1}{\beta} \sum_{\omega_n} \frac{1}{\mathcal{V}} \sum_{\mathbf{k}\sigma} \mathcal{G}_{0,\sigma}(\mathbf{k} + \mathbf{q}, i\omega_n + i\nu_\ell) \mathcal{G}_{0,\sigma}(\mathbf{k}, i\omega_n) \quad (9.164)$$

we know the form of $\mathcal{G}_{0,\sigma}$:

$$\mathcal{G}_{0,\sigma}(\mathbf{k}, z) = \frac{1}{z - \xi_{\mathbf{k}}} \quad (9.165)$$

hence our $F(z)$ has two simple poles and converges to 0 faster than $\frac{1}{z}$. Hence we can use the contour-integral method, which tells us that

$$\chi_0(\mathbf{q}, i\nu_\ell) = \frac{1}{\mathcal{V}} \sum_{\mathbf{k}\sigma} \left(\frac{f_{\text{FD}}(\xi_{\mathbf{k}+\mathbf{q}}) - f_{\text{FD}}(\xi_{\mathbf{k}})}{\xi_{\mathbf{k}+\mathbf{q}} - \xi_{\mathbf{k}} - i\nu_\ell} \right) \quad (9.166)$$

This Matsubara susceptibility is related to the regular susceptibility by an analytic continuation; $i\nu_\ell \rightarrow \omega + i\delta$, which means we have obtained quite a powerful result by not doing a lot of complicated maths.

We can do something very similar for bosons, where we use bosonic Matsubara frequencies, and we just need to be careful because there is a factor minus one that we didn't need to take into account in the fermionic case.

9.6.4 Generating Functional and Wick's Theorem

We can develop Wick's theorem in the same fashion as we developed it for zero temperature, however, the difference is that the integrals no longer go from $-\infty$ to ∞ but rather from 0 to β , and we replace $\mathcal{Z}_I[\bar{\eta}, \eta] \rightarrow \frac{Z[\bar{\eta}, \eta]}{Z_0}$. Similarly to before we will consider

$$H_0 = \sum_{\lambda} \epsilon_{\lambda} \psi_{\lambda}^{\dagger} \psi_{\lambda} \quad (9.167)$$

$$V(\tau) = - \sum_{\lambda} [\bar{\eta}_{\lambda}(\tau) \psi_{\lambda} + \psi_{\lambda}^{\dagger} \eta(\tau)] \quad (9.168)$$

plugging this into our expression for Z we obtain

$$\frac{Z[\bar{\eta}, \eta]}{Z_0} = \exp \left(- \sum_{\lambda} \int_0^{\beta} d\tau_1 d\tau_2 \bar{\eta}(\tau_1) \mathcal{G}_{\lambda}(\tau_1 - \tau_2) \eta_{\lambda}(\tau_2) \right) \quad (9.169)$$

where

$$\mathcal{G}_{\lambda}(\tau_1 - \tau_2) = - \langle T \psi_{\lambda}(\tau_1) \psi_{\lambda}^{\dagger}(\tau_2) \rangle \quad (9.170)$$

By taking various functional derivatives of our new generating functional, we can derive an imaginary-time Wick's theorem

$$\mathcal{G}(1, 2, \dots, n; n', \dots, 2', 1') = \sum_P \prod_{r=1}^n (-1)^P \mathcal{G}(r - P'_r) \quad (9.171)$$

which looks identical to the zero-temperature version, except that we use τ instead of t . Now we can also relate the full partition function the non-interacting partition function by means of a Feynman Diagram expansion;

$$Z = Z_0 \sum \{ \text{unlinked Feynman - Diagrams} \} \quad (9.172)$$

And the shift in the *Helmholtz Free Energy*

$$\Delta F = F - F_0 = - \frac{1}{\beta} \sum \{ \text{linked Feynman - Diagrams} \} \quad (9.173)$$

The Feynman rules for imaginary time diagrams are very similar to those we had in Table 9.1. However, now it is the Matsubara Green's functions that we obtain, and interaction lines are $-V(q)$ instead of $iV(q)$.

9.6.5 Electron in a Disordered Potential

In real physical systems the atoms in a lattice are not located in a perfectly periodic grid, and there are often other types of impurities, for instance a different atom, with different properties placed in the midst of the other atoms perturbing the crystal structure and characteristics.

Let us consider the following Hamiltonian:

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}} + V_{\text{disorder}}, \quad V_{\text{disorder}} = \int d^3\mathbf{x} U(\mathbf{x}) \psi^{\dagger}(\mathbf{x}) \psi(\mathbf{x}) \quad (9.174)$$

where $U(\mathbf{x})$ is the potential due to the impurities:

$$U(\mathbf{x}) = \sum_j \mathcal{U}_j(\mathbf{x} - \mathbf{R}_j) \quad (9.175)$$

where \mathbf{R}_j is the location of the j^{th} impurity. We're not interested in a specific example, where we know exactly where all the impurities are, but rather a general result, where we've averaged over all impurity positions. Hence we're interested in the impurity averaged Green's function, which we will denote by $\overline{\mathcal{G}(\mathbf{k}, \mathbf{k}', i\omega_n)}$. First let us denote

$$\delta U(x) = U(x) - \overline{U(x)} = U(x) - \mu \quad (9.176)$$

hence it is only the fluctuation in $U(x)$ that is important; the averaged value just readjusts the chemical potential. Hence we can use the following potential

$$V = \int_{\mathbf{k}\mathbf{k}'} \delta U_{\mathbf{k}-\mathbf{k}'} c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} = \begin{array}{c} \times \\ \vdots \\ \longrightarrow \quad \longrightarrow \\ \mathbf{k} \quad \mathbf{k}' \end{array} \quad (9.177)$$

Note that the impurity average of δU is zero, therefore the first order term in our expansion for $\overline{G(\mathbf{k}, \mathbf{k}', i\omega_n)}$ will be zero; we must go to second order

$$\overline{G(\mathbf{k}, \mathbf{k}', i\omega_n)} = G_0(\mathbf{k}, \mathbf{k}', i\omega_n) \delta_{\mathbf{k}, \mathbf{k}'} + G_0(\mathbf{k}) \int_{\mathbf{k}_1} G_0(\mathbf{k}_1) \overline{\delta U_{\mathbf{k}-\mathbf{k}_1} \delta U_{\mathbf{k}_1-\mathbf{k}'}} G_0(\mathbf{k}') \quad (9.178)$$

we can show that

$$\overline{\delta U_{\mathbf{k}-\mathbf{k}_1} \delta U_{\mathbf{k}_1-\mathbf{k}'}} = \delta(\mathbf{k} - \mathbf{k}') n_I |u(\mathbf{k} - \mathbf{k}_1)|^2 \quad (9.179)$$

where n_I is the density of impurities and $u(\mathbf{q})$ are the Fourier coefficients of the impurity potential. We obtain a Dyson Equation, which allows us to write the full Green's function in terms of the impurity-free Green's function and a self-energy:

$$G(\mathbf{k}, \mathbf{k}') = \frac{1}{i\omega - \xi_{\mathbf{k}} - \Sigma(\mathbf{k})} \quad (9.180)$$

where

$$\Sigma(\mathbf{k}) = n_I \int_{\mathbf{k}_1} \frac{|u(\mathbf{k} - \mathbf{k}_1)|^2}{i\omega_n - \xi_{\mathbf{k}_1}} \approx -\frac{i \text{sgn}(\omega_n)}{2\tau} \quad (9.181)$$

where we've defined the scattering rate $\tau \equiv 2\pi n_I u_0^2$ and we've assumed that the Fourier coefficients $u(\mathbf{k} - \mathbf{k}_1)$ are dominated by the constant part, u_0 . This gives us

$$G(\mathbf{k}, i\omega_n) = \frac{1}{i\omega - \xi_{\mathbf{k}} + i \frac{\text{sgn}(\omega_n)}{2\tau}} \quad (9.182)$$

which tells us that the impurity gives rise to electron-lifetimes; or equivalently the existence of impurities in metals gives rise to a resistivity, just as one might expect.

9.6.6 Electron-Phonon Coupling

Let us consider a system of phonons and electrons that coupled through an interaction given by $V_{e-\text{ph}}$:

$$H_{\text{ph}} = \sum_{\mathbf{q}\lambda} \omega_{\mathbf{q}\lambda} \left(a_{\mathbf{q}\lambda}^\dagger a_{\mathbf{q}\lambda} + \frac{1}{2} \right) \quad (9.183)$$

$$H_e = \sum_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \quad (9.184)$$

$$V_{e-\text{ph}} = \int d^3\mathbf{x} (-e\hat{\rho}(\mathbf{x})) \sum_j U_{\text{ion}}(\mathbf{x} - \mathbf{R}_j) \quad (9.185)$$

In the low-temperature limit we can expect the displacement $R_j - R_j^0$ to be sufficiently small that we can approximate the ion-potential up to first order:

$$U_{\text{ion}}(\mathbf{x} - \mathbf{R}_j) = U_{\text{ion}}(\mathbf{x} - \mathbf{R}_j^0) - \nabla_{\mathbf{x}} U_{\text{ion}}(\mathbf{x} - \mathbf{R}_j^0) \cdot \mathbf{u}_j \quad (9.186)$$

where $\mathbf{u}_j = \mathbf{R}_j - \mathbf{R}_j^0$. The first term is a periodic potential, which tells us that to zeroth order in \mathbf{u}_j the eigenfunctions are Bloch-waves. The vector \mathbf{u}_j can be written as

$$\mathbf{u}_j = \sum_{\mathbf{q}\lambda} \bar{\varepsilon}_\lambda e^{-i\mathbf{q}\cdot\mathbf{R}_j^0} \frac{\ell_{\mathbf{q}\lambda}}{\sqrt{2}} (a_{\mathbf{q}\lambda} + a_{-\mathbf{q}\lambda}^\dagger) \quad (9.187)$$

the quantity $a + a^\dagger$ is related to the displacement (recall the harmonic oscillator), and the displacement of the j^{th} atom is related to the phonon-displacement because phonons are waves in the periodic lattice structure. The quantity $\ell_{\mathbf{q}\lambda} = \sqrt{\frac{\hbar}{2mN\omega_{\mathbf{q}\lambda}}}$ is the so-called *magnetic-length*. Hence, the interaction up to first order is given by

$$V_{\text{e-ph}} = \sum_{\mathbf{q}\mathbf{k}\lambda} g_{\mathbf{q}\lambda} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} [a_{\mathbf{q}\lambda} + a_{-\mathbf{q}\lambda}^\dagger] \quad (9.188)$$

where the electron-phonon coupling constant, $g_{\mathbf{q}\lambda}$ is given by

$$g_{\mathbf{q}\lambda} = \begin{cases} Cq\ell_{\mathbf{q}\lambda} & \lambda : \text{longitudinal} \\ 0 & \lambda : \text{transversal} \end{cases} \quad (9.189)$$

at each vertex we write $(ig_{\mathbf{q}\lambda})$ as this is the coupling-constant, which means that we should expand in terms of g .

The non-interacting Green's functions are

$$\mathcal{G}^0(\mathbf{k}, i\omega_n) = \frac{1}{i\omega_n - \tilde{\epsilon}_{\mathbf{k}}} \quad (9.190)$$

$$\mathcal{D}^0(\mathbf{q}, i\nu_n) = \frac{2\omega_{\mathbf{q}}}{(i\nu_n)^2 - \omega_{\mathbf{q}}^2} \quad (9.191)$$

This interaction is quite similar to the Coulomb interaction, however, the propagation speed is lower than that of photons, which tells us that the electron-phonon interaction is always delayed. In some materials at low temperature this can lead to attractive forces between electrons, such that they form Cooper-pairs; superconductors.

If we now just consider the lowest order contribution to the self-energy:

$$\Sigma(k) = \text{---} \overset{\mathbf{q}}{\text{---}} \text{---} \quad (9.192)$$

The diagram shows a horizontal line representing an electron propagator with momentum k and energy $i\omega_\ell$. A wavy line representing a phonon propagator with momentum q and energy $i\nu_n$ is attached to this line. The electron propagator is split into two segments with momenta k and $k-q$, and the phonon propagator is labeled with q above it.

which is

$$\Sigma(\mathbf{k}, i\omega_\ell) = \sum_{\mathbf{q}, n} \frac{-g_{\mathbf{q}\lambda}^2}{\beta} \left(\frac{1}{i\omega_\ell - i\nu_n - \epsilon_{\mathbf{k}-\mathbf{q}}} \right) \left[\frac{1}{i\nu_n - \omega_{\mathbf{q}}} - \frac{1}{i\nu_n + \omega_{\mathbf{q}}} \right] \quad (9.193)$$

where I've written the phonon propagator out. Performing the contour-integral we obtain

$$\Sigma(\mathbf{k}, z) = \sum_{\mathbf{q}} g_{\mathbf{q}}^2 \left[\frac{1 + n_{\mathbf{q}} - f_{\mathbf{k}-\mathbf{q}}}{z - (\epsilon_{\mathbf{k}-\mathbf{q}} + \omega_{\mathbf{q}})} + \frac{n_{\mathbf{q}} + f_{\mathbf{k}-\mathbf{q}}}{z - (\epsilon_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{q}})} \right] \quad (9.194)$$

at zero-temperature we can neglect the Bose-Einstein distribution. The first term in the self-energy corresponds to virtual phonon emission by an electron, whereas the second term corresponds to virtual phonon emission by a hole.

The real part of this energy contributes to the renormalised energy through the self-consistent equation

$$\varepsilon_{\mathbf{k}}^* = \varepsilon_{\mathbf{k}} + \text{Re}(\Sigma(\mathbf{k}, \varepsilon_{\mathbf{k}}^*)) \quad (9.195)$$

whereas the imaginary part contributes to the decay rate of the electron, which by means of the Cauchy-Dirac relation can be shown to give

$$\Gamma_{\mathbf{k}} = 2\pi \sum_{\mathbf{q}} g_{\mathbf{q}}^2 \left[(1 + n_{\mathbf{q}} - f_{\mathbf{k}-\mathbf{q}}) \delta(\varepsilon_{\mathbf{k}} - (\varepsilon_{\mathbf{k}-\mathbf{q}} + \omega_{\mathbf{q}})) + (n_{\mathbf{q}} + f_{\mathbf{k}-\mathbf{q}}) \delta(\varepsilon_{\mathbf{k}} - (\varepsilon_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{q}})) \right] \quad (9.196)$$

which is reminiscent of something one may obtain from Fermi's Golden Rule.

9.7 FLUCTUATION-DISSIPATION THEOREM AND LINEAR RESPONSE THEORY

9.7.1 Classical Harmonic Oscillator

The differential equation that describes a classical harmonic oscillator

$$m(\ddot{x} + \omega_0^2 x) + \eta \dot{x} = f(t) \quad (9.197)$$

can be solved by means of a Fourier transform

$$x(\omega) = \chi(\omega) f(\omega), \quad \chi(\omega) = \frac{1}{m(\omega_0^2 - \omega^2) - i\omega\eta} \quad (9.198)$$

if we now denote $\chi = \chi' + i\chi''$ for $\chi', \chi'' \in \mathbb{R}$ (which we will do for this entire section) then

$$\chi''(\omega) = \frac{\omega\eta}{m^2(\omega_0^2 - \omega^2)^2 + \omega^2\eta^2} = \omega\eta |\chi(\omega)|^2 \quad (9.199)$$

In the classical example we can relate this imaginary part to the noise spectrum, just as we will do for the quantum mechanical case:

$$S(\omega) = \langle |x(\omega)|^2 \rangle = \frac{2k_B T}{\omega} \chi''(\omega) \quad (9.200)$$

9.7.2 Fluctuations and Dissipation in a Quantum World

We can derive a fluctuation-dissipation theorem in quantum mechanics by means of spectral decomposition. Let us first consider a correlation function $S(t - t') = \langle A(t) A(t') \rangle$ and consider the eigenbasis of the full Hamiltonian $H|\lambda\rangle = E_{\lambda}|\lambda\rangle$:

$$\begin{aligned} S(t - t') &= \langle A_H(t) A_H(t') \rangle = \sum_{\lambda} \langle \lambda | e^{-\beta(H-F)} A_H(t) A_H(t') | \lambda \rangle \\ &= \sum_{\lambda\lambda'} e^{-\beta(E_{\lambda}-F)} \langle \lambda | A_H(t) | \lambda' \rangle \langle \lambda' | A_H(t') | \lambda \rangle \end{aligned} \quad (9.201)$$

distributing the time-dependence from the operators to the states we get

$$S(t - t') = \sum_{\lambda\lambda'} e^{-\beta(E_{\lambda}-F)} |\langle \lambda | A | \lambda' \rangle|^2 e^{-i(E_{\lambda'}-E_{\lambda})(t-t')} \quad (9.202)$$

which we can quite easily Fourier transform, because the only time-dependence is in the exponential

$$S(\omega) = \sum_{\lambda\lambda'} e^{-\beta(E_\lambda - F)} |A_{\lambda\lambda'}|^2 2\pi\delta(E_{\lambda'} - E_\lambda - \omega) \quad (9.203)$$

Similarly, for the retarded response function $\chi_R(t - t') = i\langle [A(t), A(t')] \rangle \Theta(t - t')$ we obtain

$$\chi_R(t) = i \sum_{\lambda\lambda'} e^{\beta F} (e^{-\beta E_\lambda} - e^{-\beta E_{\lambda'}}) |A_{\lambda\lambda'}|^2 e^{-i(E_{\lambda'} - E_\lambda)t} \Theta(t) \quad (9.204)$$

and Fourier transforming

$$\chi_R(\omega) = \sum_{\lambda\lambda'} e^{\beta F} (e^{-\beta E_\lambda} - e^{-\beta E_{\lambda'}}) |A_{\lambda\lambda'}|^2 \frac{1}{(E_{\lambda'} - E_\lambda) - \omega - i\delta} \quad (9.205)$$

where we've included a convergence factor in the Fourier transformation. Denote everything except the fraction as Y . We have that $\chi_R(\omega) = \chi(\omega + i\delta)$ so if we denote $\chi(\omega + i\delta) = \chi'(\omega) + i\chi''(\omega)$ we get that

$$\chi(\omega + i\delta) = Y \left(\mathcal{P} \frac{1}{(E_{\lambda'} - E_\lambda) - \omega} + i\pi\delta[(E_{\lambda'} - E_\lambda) - \omega] \right) \quad (9.206)$$

Note that

$$\chi'(\omega) = \int \frac{d\omega'}{\pi} \mathcal{P} \frac{1}{\omega' - \omega} \chi''(\omega') \quad (9.207)$$

which tells us that

$$\chi(z) = \int \frac{d\omega'}{\pi} \frac{1}{\omega' - z} \chi''(\omega') \quad (9.208)$$

which is a *Kramers-Kronig relation*.

Using our expressions for $S(\omega)$ and $\chi''(\omega)$ we can show that

$$S(\omega) = 2\hbar(1 + n_B(\hbar\omega))\chi''(\omega) \quad (9.209)$$

CONDENSED MATTER THEORY 2

10.1 PATH INTEGRALS

In this section we consider imaginary-time Feynman path integrals, which is the first stepping stone in our development of the theory behind superconductivity, itinerant magnetism and non-equilibrium many-body physics. The idea of path integrals is to cut up a matrix element of the form $\langle f | \exp(-iHt) | i \rangle$ into small pieces and summing over all possible paths. Effectively

$$\langle f | \exp(-iHt) | i \rangle = \sum_{\text{paths}} \exp(iS_{\text{path}}) \quad (10.1)$$

where we've defined the action as

$$S_{\text{path}} \equiv \int_0^t dt' (pq - H[p, q]) \quad (10.2)$$

This is a precise formulation of quantum mechanics. This concept can readily be extended to imaginary times, such that we can treat temperature dependent problems. In the many-body problem the matrix element of interest is

$$Z = \text{tr} [\exp(-\beta H)] = \sum_{\lambda} \langle \lambda | \exp(-iHt) | \lambda \rangle \Big|_{t \rightarrow -i\beta} \quad (10.3)$$

where now, using imaginary times

$$Z = \sum_{\text{periodic paths}} \exp(-S_E), \quad (10.4)$$

and

$$S_E = \int_0^\beta d\tau (-ip\partial_\tau q + H[p, q]) \quad (10.5)$$

However, this idea can further be extended, such that we obtain a rich description of many-body physics, which we get by introducing *coherent states*. For a single boson field, a coherent state is given by

$$|b\rangle = e^{b\hat{b}^\dagger} |0\rangle \quad (10.6)$$

similarly in the many-body case the coherent state is an eigenstate of the field operator, $\hat{\psi}$:

$$\hat{\psi}(x)|\phi\rangle = \phi(x)|\phi\rangle \quad (10.7)$$

This will allow us to effectively replace the fields $\hat{\psi}$ and $\hat{\psi}^\dagger$ with their eigenvalues ϕ and $\bar{\phi}$, however, this shouldn't blindly be done – for instance we must require that a Hamiltonian is normal ordered in order for us to perform the aforementioned replacement. Note that this also implies that time-ordered expectation values such as

$$\langle T \hat{\psi}(1) \hat{\psi}^\dagger(2) \rangle = \frac{1}{Z} \sum_{\text{path}} \exp(-S_{\text{path}}) \phi(1) \bar{\phi}(2) \quad (10.8)$$

just become a functional integral over all possible paths connecting the points 1 and 2 (we no longer need to think about the operator properties of the fields).

10.1.1 Coherent States for Bosons

As defined above, the bosonic coherent state is

$$|b\rangle \equiv e^{b\hat{b}^\dagger}|0\rangle = \sum_n \frac{b^n |n\rangle}{\sqrt{n!}} \quad (10.9)$$

where I have written out the exponential. This has a few properties that are noteworthy, for instance

$$\hat{b}|b\rangle = b|b\rangle, \quad \langle b|\hat{b}^\dagger = \langle b|\bar{b} \quad (10.10)$$

Note that the overlap between two coherent states is

$$\langle b_1|b_2\rangle = e^{\bar{b}_1 b_2} \quad (10.11)$$

Generally matrix elements of *normal ordered* operators, $\hat{O}[\hat{b}^\dagger, \hat{b}]$, can be evaluated by replacing the operators with their eigenvalues, and multiplying by the overlap:

$$\langle b_1|\hat{O}[\hat{b}^\dagger, \hat{b}]|b_2\rangle = O[\bar{b}_1, b_2]e^{\bar{b}_1 b_2} \quad (10.12)$$

The completeness relation must be modified slightly, due to the overlap between the states

$$\mathbb{1} = \int \frac{d\bar{b}db}{2\pi i} e^{-\bar{b}b}|b\rangle\langle b| \quad (10.13)$$

We say that the coherent states form an *overcomplete* basis, which is similar to a complete basis, but the coefficients that describe an arbitrary state are not uniquely defined.

BOSONIC GAUSSIAN INTEGRAL

$$\int \prod_j \frac{d\bar{b}_j db_j}{2\pi i} e^{-[\bar{\mathbf{b}} \cdot \mathbf{A} \cdot \mathbf{b} - \bar{\mathbf{j}} \cdot \mathbf{b} - \bar{\mathbf{b}} \cdot \mathbf{j}]} = \frac{\exp[\bar{\mathbf{j}} \cdot \mathbf{A}^{-1} \cdot \mathbf{j}]}{\det A} \quad (10.14)$$

10.1.2 Path Integral for the Partition Function: Bosons

We can make our $\sum_{\text{periodic paths}}$ a bit more rigorous by defining it as a functional integral over all possible b and \bar{b} , we obtain the working definition of the partition functional, which will be useful for the remainder of the course:

$$Z \equiv \int \mathcal{D}[\bar{b}, b] \exp(-S[\bar{b}, b]), \quad (10.15)$$

$$S[\bar{b}, b] \equiv \int_0^\beta d\tau (\bar{b} \partial_\tau b + H[\bar{b}, b])$$

This is obtained by slicing the interval $[0, \beta]$ up into infinitesimal slices and summing over all possible paths, we begin with:

$$\text{tr} [\exp(-\beta H)] = \int \frac{d\bar{b}db}{2\pi i} e^{-\bar{b}b} \langle b | \exp(-\beta H) | b \rangle \quad (10.16)$$

even if H is normal ordered, $\exp(H)$ is *not* normal ordered. So we will need to use that

$$\exp(-\beta H) = \lim_{N \rightarrow \infty} (1 - \delta\tau H)^N \quad (10.17)$$

where $\delta\tau = \frac{\beta}{N}$, then we can insert a (over-)completeness relation inbetween each of the terms:

$$\int \prod_j \frac{d\bar{b}_j db_j}{2\pi i} e^{-\bar{b}_j b_j} \langle b_j | \exp(-\delta\tau H) | b_{j-1} \rangle \quad (10.18)$$

where we've used that $b_0 = b_N$. Now the linearised exponential together with the normalisation factor gives us

$$\begin{aligned} e^{-\bar{b}_j b_j} \langle b_j | \exp(-\delta\tau H) | b_{j-1} \rangle &\approx e^{-\bar{b}_j b_j} \langle b_j | 1 - \delta\tau H | b_{j-1} \rangle \approx e^{-\bar{b}_j b_j + \bar{b}_j b_{j-1}} (1 - \delta\tau H[\bar{b}_j, b_{j-1}]) \\ &\approx \exp \left[-\delta\tau \left(b_j \frac{b_j - b_{j-1}}{\delta\tau} + H[\bar{b}_j, b_{j-1}] \right) \right] \end{aligned} \quad (10.19)$$

hence, in the continuum limit we obtain Equation 10.15. For multiple bosons we obtain a similar expression

BOSONIC PARTITION FUNCTION

$$Z = \int \mathcal{D}[\bar{\mathbf{b}}, \mathbf{b}] e^{-S}, \quad S[\bar{\mathbf{b}}, \mathbf{b}] \equiv \int_0^\beta \left(\sum_\lambda \bar{b}_\lambda \partial_\tau b_\lambda + H[\bar{b}_\lambda, b_\lambda] \right) \quad (10.20)$$

where $\mathcal{D}[\bar{\mathbf{b}}, \mathbf{b}] = \prod_\lambda \mathcal{D}[\bar{b}_\lambda, b_\lambda]$. We can use these quantities to calculate time-ordered expectation values

$$G(2-1) = -\langle T \hat{b}(2) \hat{b}^\dagger(1) \rangle = -\frac{\int \mathcal{D}[\bar{\mathbf{b}}, \mathbf{b}] e^{-S} b(2) \bar{b}(1)}{\int \mathcal{D}[\bar{\mathbf{b}}, \mathbf{b}] e^{-S}} \quad (10.21)$$

During this course we are particularly interested in *Gaussian Path Integrals*, where the action is a quadratic functional of the fields. For a free boson, with Hamiltonian $\hat{H} = \hat{b}_\alpha^\dagger h_{\alpha\beta} \hat{b}_\beta$ the partition function can be evaluated exactly:

$$Z_G = \int \mathcal{D}[\bar{b}, b] \exp \left[-\int_0^\beta d\tau \bar{b}(\partial_\tau + \underline{h}) b \right] = [\det(\partial_\tau + \underline{h})]^{-1} \quad (10.22)$$

The matrix $\partial_\tau + \underline{h}$ is better thought of when it is expressed in Matsubara space, where $b(\tau) = \beta^{-\frac{1}{2}} \sum_n b(i\nu_n) e^{-i\nu_n \tau}$ and we get

$$(\partial_\tau + h_{\alpha\beta}) \rightarrow (-i\nu_n \delta_{\alpha\beta} + h_{\alpha\beta}) \quad (10.23)$$

in which case we get

$$Z_G = \frac{1}{\det[\partial_\tau + \underline{h}]} = \prod_{n,\lambda} \frac{1}{(-i\nu_n + \epsilon_\lambda)} \quad (10.24)$$

where we have written the matrix out in Matsubara space, and decomposed the Hamiltonian to its eigenvalues, ϵ_λ . Note that this means

$$F_G = -T \ln Z_G = T \sum_{n,\lambda} \ln(\epsilon_\lambda - i\nu_n) e^{i\nu_n 0^+} \quad (10.25)$$

which is the free energy of non-interacting bosons. Generally the matrix that stands in place of $(\partial_\tau + \underline{h})$ is closely related to the Green's matrix:

$$Z_G = \int \mathcal{D}[\bar{b}, b] \exp \left[-\int_0^\beta \bar{b}(-G^{-1}) b \right] = [\det(-G^{-1})]^{-1} \quad (10.26)$$

so in general

$$F = T \ln \det [-G^{-1}] = T \text{tr} \ln [-G^{-1}] \quad (10.27)$$

where the equality $\ln \det = \text{tr} \ln$ will be useful throughout the course. We can generalise our results such that they include source terms, so that we obtain the generating functional that was used throughout CMT1:

$$Z_G[\bar{j}, j] = \int \mathcal{D}[\bar{b}, b] e^{-\int_0^\beta d1 [\bar{b}(\partial_\tau + \underline{h}) b - j(1) \cdot b(1) - \bar{b}(1) \cdot j(1)]} \frac{e^{-\int_0^\beta d1 d2 \bar{j}(1) G(1-2) j(2)}}{\det[\partial_\tau + \underline{h}]} \quad (10.28)$$

10.1.3 Fermions: Coherent States and Grassmann mathematics

The treatment for fermions follows the bosonic analysis quite closely, except that there is a fundamental difference in how we treat the eigenvalues of the field operators. For bosons we replaced $\hat{\psi} \rightarrow \phi$ and treated ϕ as a number. However, due to the fact that fermionic operators *anti-commute* we need their eigenvalues to commute too, which inspires us to use Grassmann numbers.

Similarly to before we define a coherent state as

$$|c\rangle = \exp(\hat{c}^\dagger c) |0\rangle, \quad \langle c| = \langle 0| \exp(\bar{c} \hat{c}) \quad (10.29)$$

where the numbers $c, \bar{c} \in \mathcal{G}$ *anti-commute* with themselves and with the field operators:

$$c\bar{c} + \bar{c}c = 0, \quad c\hat{\psi} + \hat{\psi}c = 0 \quad (10.30)$$

this also implies that

$$c^2 = 0, \quad \bar{c}^2 = 0 \quad (10.31)$$

(due to the fact that $(\hat{\psi}^{(+)})^2 |\Psi\rangle = |0\rangle$). This tells us that Taylor expansion of functions are almost trivial:

$$f(\bar{c}, c) = f_0 + \bar{c}f_1 + \bar{f}_1 c + f_2 \bar{c}c \quad (10.32)$$

where $f_0, f_1, \bar{f}_1, f_2 \in \mathbb{C}$. This is the full expansion; higher order terms are zero. From here we are ready to use the results from the bosonic case,

$$\sum_{\bar{b}b} = \int \frac{d\bar{b}db}{2\pi i} e^{-\bar{b}b} \rightarrow \sum_{\bar{c}c} = \int d\bar{c}dc e^{-\bar{c}c} \quad (10.33)$$

note the lack of the $2\pi i$ in the denominator for fermions. Trace formulae contain an extra minus:

$$\text{tr} A = \sum_{\bar{c}c} \langle -\bar{c} | A | c \rangle \quad (10.34)$$

and Jacobians as well as Gaussian integrals are the inverses of their bosonic counterparts:

FERMIONIC GAUSSIAN INTEGRAL

$$\int \prod_j d\bar{c}_j dc_j e^{-[\bar{c} \cdot A \cdot c - \bar{j} \cdot c - \bar{c} \cdot j]} = \det A \exp [\bar{j} \cdot A^{-1} \cdot j] \quad (10.35)$$

notice how it is no longer the inverse of the determinant that appears. The completeness relation is rather simple

$$\int d\bar{c}dc |c\rangle \langle \bar{c}| e^{-\bar{c}c} = |0\rangle \langle 0| + |1\rangle \langle 1| \equiv \mathbb{1} \quad (10.36)$$

The extension from one fermion to many fermions is as simple it was for bosons, but keep in mind that all Grassmann numbers anti-commute with each other.

IO.I.4 Path Integral for the Partition Function: Fermions

We once again begin with

$$Z = - \int d\bar{c}_N dc_0 e^{\bar{c}_N c_0} \langle \bar{c}_N | \exp(-\beta H) | c_0 \rangle \quad (10.37)$$

and use that $\exp(-\beta H) = \lim_{N \rightarrow \infty} (1 - \delta \tau H)^N$, so that

$$Z = - \int d\bar{c}_N dc_0 e^{\bar{c}_N c_0} \prod_{j=1}^{N-1} d\bar{c}_j dc_j e^{-\bar{c}_j c_j} \prod_{j=1}^N \langle \bar{c}_j | \exp(-\delta \tau H) | c_{j-1} \rangle \quad (10.38)$$

previously we used that $b_N = b_0$, but for fermions $c_N = -c_0$. Following similar steps as we did for the bosonic case we obtain

$$S = \int_0^\infty d\tau [\bar{c}(\partial_\tau + \epsilon)c], \quad Z = \int \mathcal{D}[\bar{c}, c] \exp[-S] \quad (10.39)$$

where $\mathcal{D}[\bar{c}, c] = \prod_j d\bar{c}_j dc_j$. We can transform to Matsubara space, where we now use fermionic frequencies, ω_n . Obtaining

$$S = \sum_n \bar{c}_n (-i\omega_n + \epsilon) c_n \quad (10.40)$$

IO.I.5 The Hubbard-Stratonovich Transformation

The Hubbard-Stratonovich transformation is a transformation in the action, where we introduce new fields such that we can complete the square and simplify the expression to obtain a palatable effective action.

Consider a fermionic path integral of the form

$$Z_0 = \int \mathcal{D}[\bar{\mathbf{c}}, \mathbf{c}] \exp \left[- \int_0^\beta d\tau (\bar{\mathbf{c}}(\partial_\tau + \underline{h})\mathbf{c} - g\bar{\mathbf{A}} \cdot \mathbf{A}) \right] \quad (10.41)$$

where \mathbf{A} is an electron bilinear. Now let us multiply this by a *number*

$$Z_\alpha = \int \mathcal{D}[\bar{\alpha}, \alpha] \exp \left[- \sum_j \int_0^\beta g^{-1} \bar{\alpha} \cdot \alpha \right] \quad (10.42)$$

This is just a Gaussian integral which we can evaluate exactly. Let us consider $Z = Z_0 \times Z_\alpha$, giving an effective action:

$$S = \int_0^\beta d\tau [\bar{\mathbf{c}}(-\underline{G}_0^{-1})\mathbf{c} - g\bar{\mathbf{A}} \cdot \mathbf{A} - g^{-1}\bar{\alpha} \cdot \alpha] \quad (10.43)$$

where $S = S_0 + S_\alpha$. The partition function is given by a functional integral of the action above, so we can shift the variable α freely, just like one would do for regular integrals. If we choose to shift it such that the new integration “variable” is

$$\Delta = \alpha - g\mathbf{A} \quad (10.44)$$

we get that

$$-g\bar{\mathbf{A}} \cdot \mathbf{A} - g^{-1}\bar{\alpha} \cdot \alpha = \bar{\mathbf{A}} \cdot \Delta + \bar{\Delta} \cdot \mathbf{A} + g^{-1}\bar{\Delta} \cdot \Delta \quad (10.45)$$

hence we obtain

$$Z = \int \mathcal{D}[\Delta, \mathbf{c}] \exp[-S[\Delta, \mathbf{c}]], \quad S[\Delta, \mathbf{c}] = \int_0^\beta d\tau g^{-1} \bar{\Delta} \cdot \Delta + \bar{\mathbf{c}} (-\underline{G}_0^{-1}) \mathbf{c} + \bar{\mathbf{A}} \cdot \Delta + \bar{\Delta} \cdot \mathbf{A} \quad (10.46)$$

Notice how all terms either don't depend on \mathbf{c} or are bilinears, which implies that this is an action that can be integrated by means of our Gaussian integration technique described above. However, now once the electrons are integrated out we obtain an effective action in the Δ fields.

The process of performing the \mathbf{c} integration is referred to as integrating the electrons out.

$$S_E[\bar{\Delta}, \Delta] = \int d\tau g^{-1} \bar{\Delta} \cdot \Delta - \text{tr} \ln[\partial_\tau + \underline{h}_E[\bar{\Delta}, \Delta]] \quad (10.47)$$

Coulomb Force

In the case of the three-dimensional Coulomb interaction the interacting-Hamiltonian in Fourier space is

$$H_I = \frac{1}{2} \int_{\mathbf{q}} \frac{e^2 \hat{\rho}_{\mathbf{q}} \hat{\rho}_{-\mathbf{q}}}{\epsilon_0 q^2} \quad (10.48)$$

where $\hat{\rho}_{\mathbf{q}} = \hat{c}_{\mathbf{q}}^\dagger \hat{c}_{\mathbf{q}}$. We see that this part of the Hamiltonian is quartic in the fields, so it is suitable to be integrated away by means of the Hubbard-Stratonovich transformation. Effectively we can add a bilinear in a new field $\phi_{\mathbf{q}}$ to the Hamiltonian and then perform the change of variables. This corresponds exactly to what was done above when we multiplied the partition function by a Gaussian integral. Let us choose

$$H_I \rightarrow H'_I = \frac{1}{2} \int_{\mathbf{q}} \left[\frac{e^2 \rho_{\mathbf{q}} \rho_{-\mathbf{q}}}{\epsilon_0 q^2} - \epsilon_0 q^2 \phi_{\mathbf{q}} \phi_{-\mathbf{q}} \right] \quad (10.49)$$

and now we let

$$\phi_{\mathbf{q}} \rightarrow \phi_{\mathbf{q}} - \frac{e \rho_{\mathbf{q}}}{\epsilon_0 q^2} \quad (10.50)$$

to obtain

$$H'_I = \int_{\mathbf{q}} \left(e \rho_{\mathbf{q}} \phi_{-\mathbf{q}} - \frac{\epsilon_0}{2} q^2 \phi_{\mathbf{q}} \phi_{-\mathbf{q}} \right) \quad (10.51)$$

which is a bilinear in both fields and hence can be handled more easily.

10.2 PATH INTEGRALS AND ITINERANT MAGNETISM

One application of the Hubbard-Stratonovich transformation is in describing itinerant magnetism. The Weiss field, \mathbf{M} , which is more commonly known as the magnetisation, is a quantity one normally encounters in a mean-field-theory discussion of magnetism. However, this field appears in the path integral when we perform a Hubbard-Stratonovich transformation, as we will see shortly. Let us begin with the Hubbard model

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + U \sum_j n_{j\uparrow} n_{j\downarrow} \quad (10.52)$$

which describes electrons on a tight-binding lattice. Notice how the electron-electron interaction can be written as

$$Un_{j\uparrow}n_{j\downarrow} = \frac{U}{2} \left[(n_{j\uparrow} + n_{j\downarrow}) - (n_{j\uparrow} - n_{j\downarrow})^2 \right] \quad (10.53)$$

the first term is just a shift in the chemical potential so we will omit it, however the second term is related to the spin along the z-axis. Due to the isotropy of the system we can express this as a third of the average spin squared in all three directions:

$$(n_{j\uparrow} - n_{j\downarrow})^2 = \frac{1}{3} |\mathbf{S}_j|^2, \quad \mathbf{S}_j \equiv \sum_{\alpha\beta} c_{j\alpha}^\dagger \boldsymbol{\sigma}_{\alpha\beta} c_{j\beta} \quad (10.54)$$

so

$$Un_{j\uparrow}n_{j\downarrow} \rightarrow I |\mathbf{S}_j|^2, \quad 3I = U \quad (10.55)$$

So our action is

$$S = \int_0^\beta d\tau \left[\sum_{\mathbf{k}\sigma} \bar{c}_{\mathbf{k}\sigma} (\partial_\tau + \epsilon_{\mathbf{k}}) c_{\mathbf{k}\sigma} - \frac{I}{2} \sum_j |\mathbf{S}_j|^2 \right] \quad (10.56)$$

now we introduce the Hubbard-Stratonovich field

$$-\frac{I}{2} \sum_j |\mathbf{S}_j|^2 \rightarrow -\frac{I}{2} \sum_j |\mathbf{S}_j|^2 + \sum_j \frac{|\mathbf{m}_j|^2}{2I} \quad (10.57)$$

and shift \mathbf{m} such that it absorbs the quartic term

$$\mathbf{m}_j = \mathbf{M}_j - I \mathbf{S}_j \quad (10.58)$$

leaving

$$S = \int_0^\beta d\tau \left[\sum_{\mathbf{k}\sigma} \bar{c}_{\mathbf{k}\sigma} (\partial_\tau + \epsilon_{\mathbf{k}}) c_{\mathbf{k}\sigma} - \sum_j \left[\mathbf{M}_j \cdot \mathbf{S}_j - \frac{|\mathbf{M}_j|^2}{2I} \right] \right] \quad (10.59)$$

which is a bilinear in the electron fields, so we are ready to integrate out the electrons:

$$Z = \int \mathcal{D}[\mathbf{M}] e^{-S_E[\mathbf{M}]}, \quad e^{-S_E[\mathbf{M}]} = \int \mathcal{D}[\bar{c}, c] e^{-S[\bar{c}, c, \mathbf{M}]} \quad (10.60)$$

where once we integrate out the electrons

$$e^{-S_E[\mathbf{M}]} = \det [\partial_\tau + h_E[\mathbf{M}]] \exp \left[- \sum_j \int_0^\beta d\tau \frac{|\mathbf{M}_j|^2}{2I} \right], \quad h_E[\mathbf{M}] \equiv \epsilon_{\mathbf{k}} \delta_{\mathbf{k}, \mathbf{k}'} - \mathbf{M}_{\mathbf{k}-\mathbf{k}'} \cdot \boldsymbol{\sigma} \quad (10.61)$$

If we write $G_0 = (i\omega - \epsilon_{\mathbf{k}})^{-1}$ and $V_{k,k'} = \mathbf{M}_{k-k'} \cdot \boldsymbol{\sigma}$ then the effective action can be expressed as

$$S_E[\mathbf{M}] = -\text{tr} \ln [-G_0^{-1}(k)] - \text{tr} \ln [1 - G_0(k) V_{k,k'}] + N\beta \sum_q \frac{|\mathbf{M}_q|^2}{2I} \quad (10.62)$$

where k and q are Matsubara-four-momenta. The second term can be expanded

$$\text{tr} \ln(1 - G_0 V) = \text{tr} \left[-G_0 V - \frac{1}{2} (G_0 V)^2 - \frac{1}{3} (G_0 V)^3 + \dots \right] \quad (10.63)$$

which can be written in terms of Feynman diagrams where an electron scatters off the Weiss field n times (for the n -th term). Hence we obtain an expression for the free energy in terms of an infinite sum of Feynman Diagrams, just like we did in CMT1, however, this time by means of a very different method.

10.2.1 *Saddle Points and the Mean-Field Theory of the Magnetism*

In the saddle-point approximation we assume that the partition is equal to its value at the saddle point, for instance where $\mathbf{M} = \mathbf{M}_0$, where

$$\left. \frac{\delta S_E[\mathbf{M}]}{\delta \mathbf{M}} \right|_{\mathbf{M}=\mathbf{M}_0} = 0 \quad (10.64)$$

we assume that

$$Z = \int \mathcal{D}[\mathbf{M}] \exp[-S_E[\mathbf{M}]] \approx \exp[-S_E[\mathbf{M}_0]] \quad (10.65)$$

let us evaluate Equation 10.64 using Equation 10.60:

$$\frac{\delta S_E}{\delta \mathbf{M}_j} = e^{S_E} \int \mathcal{D}[\bar{c}, c] \left(\frac{\mathbf{M}_j}{I} - \bar{c}_j \sigma c_j \right) e^{-S[\bar{c}, c, \mathbf{M}]} \quad (10.66)$$

requiring that this is zero means

$$\mathbf{M}_j^0 = I \langle c_j^\dagger \sigma c_j \rangle \quad (10.67)$$

which is exactly what the mean-field-theory approach would have begun with. We're often interested in a static saddle plot, such that \mathbf{M}_j^0 doesn't depend on τ , in which case

$$e^{-S_E[\mathbf{M}^0]} = \text{tr} \left[e^{-\beta \hat{H}_{\text{MF}}} \right], \quad \hat{H}_{\text{MF}} = c^\dagger h_E[\mathbf{M}^0] c + \sum_j \frac{|\mathbf{M}_j^0|^2}{2I} \quad (10.68)$$

Simplifying things further by assuming that we are describing a ferromagnet, which has a uniform magnetisation $\mathbf{M}_j^0 = M \hat{\mathbf{z}}$, in which case we get

$$H_{\text{MF}} = \sum_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}^\dagger (\epsilon_{\mathbf{k}} - \sigma_z M) c_{\mathbf{k}\sigma} + N \frac{M^2}{2I} \quad (10.69)$$

which implies that

$$S_E = - \sum_{\mathbf{k}, i\omega_n} \text{tr} \ln [\epsilon_{\mathbf{k}} - M \sigma_z - i\omega_n] + N \beta \frac{M^2}{2I} \quad (10.70)$$

from which we can obtain information about the macroscopic system. Performing the Matsubara sum:

$$F_E[M] = -\frac{1}{N\beta} \sum_{\mathbf{k}, \sigma=\pm 1} \ln [1 + e^{-\beta(\epsilon_{\mathbf{k}} - \sigma M)}] + \frac{M^2}{2I} \quad (10.71)$$

For instance by evaluating the free energy and minimising it we can obtain Stoner's criterion for ferromagnetic order.

odd terms die due to the σ sum.

Let us use this result to derive the Landau expansion of the free energy. We have

$$-\frac{T}{2} \sum_{\sigma=\pm 1} \ln [1 + e^{-\beta(\epsilon - \sigma M)}] = -T \ln [1 + e^{-\beta\epsilon}] + \sum_{r=1}^{\infty} \frac{M^{2r}}{(2r)!} \frac{d}{d\epsilon} [r-1] f(\epsilon) \epsilon \quad (10.72)$$

So

$$F[M] = F_0 + \sum_{r=1}^{\infty} \frac{M^{2r}}{(2r)!} \int d\epsilon N(\epsilon) \frac{d}{d\epsilon} [r-1] f(\epsilon) \epsilon + \frac{M^2}{2I} \quad (10.73)$$

integrating by parts so that we only have one derivative of the Fermi-Dirac function we obtain, up to fourth order in M :

$$F = F_0 + \frac{1}{2}M^2 (I^{-1} - \chi_0(T)) + \frac{M^4}{4!} \overline{(-N''(0))} + O(M^6) \quad (10.74)$$

where

$$\chi_0(T) \equiv - \int d\epsilon N(\epsilon) \frac{\partial f(\epsilon)}{\partial \epsilon} \quad (10.75)$$

and $\overline{N''(0)}$ denotes the thermal average of the second derivative of the density of states around the Fermi energy.

10.2.2 Quantum Fluctuations in the Magnetisation

Let us now consider a system that is *close* to the saddle point, that is

$$\mathbf{M}_j(\tau) = \mathbf{M}^0 + \delta \mathbf{M}_j(\tau) \Leftrightarrow \mathbf{M}_q = \mathbf{M}^0 \delta_{q=0} + \delta \mathbf{M}_q \quad (10.76)$$

where $q = (\mathbf{q}, i\nu_n)$. Near the saddle point

$$S_E[\mathbf{M}] = S_E[\mathbf{M}^0] + \frac{1}{2} \sum_{q,m,n} \frac{\delta^2 S}{\delta M_q^m \delta M_{-q}^n} \delta \mathbf{M}_q^m \delta \mathbf{M}_{-q}^n + O(\delta \mathbf{M}^3) \quad (10.77)$$

The linear term disappears because we are at a saddle point. Provided the fluctuations are small compared to the order parameter we can use this quadratic approximation to describe magnetic fluctuations.

The renormalised Green's function before we've taken the magnetisation fluctuations into account is

$$\Rightarrow = \underline{G}_k = (i\omega_n - \epsilon_{\mathbf{k}} - \sigma_z M)^{-1} \quad (10.78)$$

Now if we expand the effective action by substituting $\mathbf{M}_{k-k'} = \mathbf{M} \delta_{k-k'} + \delta \mathbf{M}_{k-k'}$:

$$F_E[\mathbf{M}] = -\frac{1}{N\beta} \text{tr} \ln \left(-\underline{G}_k^{-1} \delta_{k-k'} - \delta \mathbf{M}_{k-k'} \cdot \boldsymbol{\sigma} \right) + \sum_q \frac{|M \hat{\mathbf{z}} \delta_q + \delta \mathbf{M}_q|^2}{2I} \quad (10.79)$$

We can write this out as a sum of Feynman Diagrams, just like in Equation 10.63. However, only looking at the *change* in the free energy will remove one of the terms, which leaves us with

$$\Delta F_E[\mathbf{M}] = - \left[\text{diagram 1} + \text{diagram 2} + \dots \right] + \sum_q \frac{|\delta \mathbf{M}_q|^2}{2I} \quad (10.80)$$

so if we just include the quadratic terms (remember each vertex carries weight $\sim \delta \mathbf{M}$) we obtain

$$\Delta F_G[\mathbf{M}] = \frac{1}{2} \sum_q \delta \mathbf{M}_q \left[\frac{1}{I} - \chi_q^{(0)} \right] \delta \mathbf{M}_{-q} \quad (10.81)$$

where the bare susceptibility of the polarised metal is defined by

$$\chi_{mn}^{(0)}(q) = -\frac{1}{\beta N} \sum_k \text{tr} \left[\sigma_m \underline{G}_{k+q} \sigma_n \underline{G}_k \right] \quad (10.82)$$

the presence of the σ_z term in the Green's function means that χ contains off diagonal components (if $\underline{G} \sim \mathbb{1}$ then we get a δ_{mn}). It is convenient to change basis, so that we are working in the raising and lowering basis of the transverse spin:

$$M_q^\pm = M_q^x \pm iM_q^y, \quad \sigma_\pm = \frac{1}{2}(\sigma_x \pm i\sigma_y) \quad (10.83)$$

where now the nonzero components of the transverse susceptibility are

$$\chi_{\pm\mp}^{(0)}(q) = -\frac{1}{\beta N} \sum_k \text{tr} \left[\sigma_\pm \underline{G}_{k+q} \sigma_\mp \underline{G}_k \right] \quad (10.84)$$

note how $\chi_{-+}(q) = \chi_{+-}(-q)$ which can be shown by shifting the k -sum by $k \rightarrow k + q$ and using the cyclicity of the trace. Now writing $\delta \mathbf{M} \cdot \boldsymbol{\sigma} = \delta M^z \sigma_z + \delta M^+ \sigma_- + \delta M^- \sigma_+$, and using that the magnetisation is a real variable implying that $\delta \bar{M}_q^\pm = \delta M_{-q}^\mp$ we obtain

$$\begin{aligned} \Delta F_G[\mathbf{M}] = & \frac{1}{2} \sum_q \left[\delta M_{-q}^z \left(I^{-1} \mathbb{1} + \chi_{zz}^{(0)} \right) \delta M_q^z + \delta \bar{M}_q^+ \left((2I)^{-1} \mathbb{1} + \chi_{+-}^{(0)}(q) \right) \delta M_q^+ \right. \\ & \left. + \delta \bar{M}_{-q}^+ \left((2I)^{-1} \mathbb{1} + \chi_{-+}^{(0)}(q) \right) \delta M_{-q}^+ \right] \end{aligned} \quad (10.85)$$

which simplifies to

$$\Delta F_G[\mathbf{M}] = \sum_q \left[\frac{1}{2} \delta M_{-q}^z \left(I^{-1} \mathbb{1} + \chi_{zz}^{(0)} \right) \delta M_q^z + \delta \bar{M}_q^+ \left((2I)^{-1} \mathbb{1} + \chi_{+-}^{(0)}(q) \right) \delta M_q^+ \right] \quad (10.86)$$

when we use $\delta \bar{M}_q^\pm = \delta M_{-q}^\mp$. So we have obtained a distribution function for the Gaussian magnetic fluctuations about the Stoner mean-field theory for an itinerant ferromagnet:

$$p[\mathbf{M}] \sim e^{-\Delta S[\mathbf{M}]} = e^{-\beta N \Delta F_G[\mathbf{M}]} \quad (10.87)$$

From this Gaussian form we can directly read off the fluctuations in the magnetisation

$$\langle \delta M_q^\alpha \delta M_{-q'}^\beta \rangle = \frac{1}{\beta N} \delta_{qq'} \langle \delta M_q^\alpha \delta M_{-q}^\beta \rangle \quad (10.88)$$

where

$$\langle \delta M_q^z \delta M_{-q}^z \rangle = \frac{1}{I^{-1} - \chi_{zz}^{(0)}(q)}, \quad \langle \delta M_q^+ \delta \bar{M}_{-q}^+ \rangle = \frac{1}{(2I)^{-1} - \chi_{+-}^{(0)}(q)} \quad (10.89)$$

which we can relate to spin-correlation functions by referring back to the Hubbard-Stratonovich transformation we performed (Equation 10.58):

$$IS_j^\alpha = M_j^\alpha - m_j^\alpha \rightsquigarrow I^2 \langle S_j^\alpha S_j^\beta \rangle = \langle M_j^\alpha M_j^\beta \rangle + \langle m_j^\alpha m_j^\beta \rangle \quad (10.90)$$

the cross-terms $\langle M^\alpha m^\beta \rangle = 0$ because m and M are independent variables. We have that $\langle m_q^\alpha m_{-q}^\beta \rangle = I \delta_{\alpha\beta}$ and hence

LONGITUDINAL	TRANSVERSAL
$\langle \sigma_q^z \sigma_{-q}^z \rangle = \frac{\chi_{zz}^{(0)}(q)}{1 - I \chi_{zz}^{(0)}(q)}$	$\langle \sigma_q^z \sigma_{-q}^q \rangle = \frac{\chi_{+-}^{(0)}(q)}{1 - I \chi_{+-}^{(0)}(q)}$

these are the *RPA spin fluctuations* of itinerant ferromagnetism.

When one performs a Legendre transformation to go from a Lagrangian to a Hamiltonian the variables one introduces are also independent of initial variables

IO.2.3 Linhard Function

We can evaluate the $\chi_{+-}(q)$ exactly in the case that $\epsilon_k = \frac{k^2}{2m} - \mu$.

$$\chi_{+-}^{(0)}(q) = -\frac{1}{N\beta} \sum_{\mathbf{k}, i\omega_n} \text{tr} \left[\sigma_+ \underline{G}_{\mathbf{k}+q} \sigma_- \underline{G}_{\mathbf{k}} \right] = -\frac{1}{N\beta} \sum_{\mathbf{k}, i\omega_n} \text{tr} \left[\underline{G}_{\downarrow}(\mathbf{k}+q) \underline{G}_{\uparrow}(\mathbf{k}) \right] \quad (10.91)$$

$$= a^3 \int_{\mathbf{k}} \frac{f_{\mathbf{k}\uparrow} - f_{\mathbf{k}+\mathbf{q}\downarrow}}{(\epsilon_{\mathbf{k}+\mathbf{q}\downarrow} - \epsilon_{\mathbf{k}\uparrow}) - i\nu_n}$$

where $\epsilon_{\mathbf{k}\sigma} = \epsilon_{\mathbf{k}} - \text{sgn}(\sigma)M$, which we can evaluate exactly by expanding this into the sum of terms and changing variables such that only $f_{\mathbf{k}\sigma}$ appears. We get

$$\chi_{+-}^{(0)}(\mathbf{q}, \nu) = \frac{1}{2} \sum_{\sigma} \left(\frac{mk_{F\sigma}}{\pi^2} \right) \mathcal{F} \left(\frac{q}{2k_{F\sigma}}, \sigma \left(\frac{\nu - 2m}{4\epsilon_F} \right) \right) \quad (10.92)$$

where $k_{F\sigma} = k_F (1 + \text{sgn}(\sigma)M\epsilon_F^{-1})^{\frac{1}{2}}$ and the Lindhard function is

$$\mathcal{F}(x, y) = \frac{1}{8x} \left[(1 - A^2) \ln \left(\frac{A+1}{A-1} \right) + 2A \right], \quad A \equiv x - x^{-1}y \quad (10.93)$$

Coleman does not have the σ in his definition of $k_{F\sigma}$

IO.3 SUPERCONDUCTIVITY AND BCS THEORY

In this section we will discuss BCS Theory and how it relates to Ginzburg-Landau theory. But first we need to show that phonon-mediation can lead to an *attractive* electron-electron interaction

IO.3.1 Bardeen-Pines

When describing superconductors it is important to remember that it is not only the electrons interacting with each other, but there is also a positive ionic background that contributes. The ionic background is more inert than the electrons – the proton to electron mass ratio is ~ 1000 , which means that the positive ions move far slower than electrons do. Define the ionic plasma frequency as

$$\Omega_p^2 = \frac{Ze^2n}{\epsilon_0 M} \quad (10.94)$$

which is thousands of times smaller than the electronic plasma frequency. The ionic plasma frequency sets the length scale of charge fluctuations in the ionic background.

The polarisability not only contains an electronic contribution, χ_0 , but also an ionic contribution χ_{ion} , so the RPA effective interaction is of the form

$$V_{\text{eff}} = \frac{1}{N} \frac{V(q)}{1 + V(q) [\chi_0(q) + \chi_{\text{ion}}(q)]} = \frac{V(q)}{\epsilon(q)} \quad (10.95)$$

where $\chi_0 + \chi_{\text{ion}}$ are the electronic and ionic polarisability bubbles. In the low-frequency, large $\frac{M}{m}$ limit we have

$$\epsilon(q) = 1 + \frac{\kappa^2}{q^2} - \frac{\Omega_p^2}{\nu^2} \quad (10.96)$$

which means

$$V_{\text{eff}}(\mathbf{q}, \nu) = \frac{e^2}{\epsilon_0(q^2 + \kappa^2)} \left(1 + \frac{\omega_q^2}{\nu^2 - \omega_q^2} \right), \quad \omega_q^2 = \frac{\Omega_P^2 q^2}{q^2 + \kappa^2} \quad (10.97)$$

the first term is the effective potential due to the electrons, which is frequency independent. The second term on the other hand is the frequency-dependent retarded electron-phonon interaction, which under the right circumstances can become negative – resulting in an attractive effective electron-electron interaction.

A more physical picture is given below

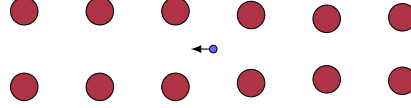


Figure 10.1: The ions behind the electron have a slightly higher density because they have been pulled inwards, thus other electrons are attracted to a point behind the original electron (which is why it's considered retarded)

10.3.2 The Cooper Instability

Leon Cooper showed that an arbitrarily small attractive force between electrons just above the Fermi sea gives rise to a two-particle bound state and a destabilisation of the Fermi-surface. Cooper imagined a state

$$|\psi(\mathbf{p})\rangle = \Lambda^\dagger(\mathbf{p})|FS\rangle, \quad \Lambda^\dagger(\mathbf{p}) \equiv \int d\mathbf{x} d\mathbf{x}' \phi_{\mathbf{x}-\mathbf{x}'} \psi_{\downarrow\mathbf{x}}^\dagger \psi_{\uparrow\mathbf{x}'}^\dagger e^{-i\mathbf{p}\cdot(\mathbf{x}+\mathbf{x}')/2} \quad (10.98)$$

Λ^\dagger is the Cooper pair creation operator, which we can rewrite in terms of the momentum-dependent field operators

$$\Lambda_{\mathbf{p}}^\dagger = \sum_{\mathbf{k}} \phi_{\mathbf{k}} c_{\mathbf{k}+\mathbf{p}/2\uparrow}^\dagger c_{-\mathbf{k}+\mathbf{p}/2\downarrow}^\dagger \quad (10.99)$$

Instead of expressing $|\psi(\mathbf{p})\rangle$ in terms of the Fermi-surface we can define a state $|\mathbf{k}, \mathbf{p}\rangle \equiv c_{\mathbf{k}+\mathbf{p}/2\uparrow}^\dagger c_{-\mathbf{k}+\mathbf{p}/2\downarrow}^\dagger |FS\rangle$ in which case we simply have

$$|\psi(\mathbf{p})\rangle = \sum_{\mathbf{k}} \phi_{\mathbf{k}} |\mathbf{k}, \mathbf{p}\rangle \quad (10.100)$$

If we now consider a Hamiltonian of the form

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + V \quad (10.101)$$

where V contains the details about the electron-electron interactions. If we assume that $|\psi(\mathbf{p})\rangle$ is an eigenstate that $H|\psi(\mathbf{p})\rangle = E_{\mathbf{p}}|\psi(\mathbf{p})\rangle$ and we obtain

$$E_{\mathbf{p}} \sum_{\mathbf{k}} \phi_{\mathbf{k}} |\mathbf{k}, \mathbf{p}\rangle = \sum_{|\mathbf{k}\pm\mathbf{p}/2| > k_F} (\epsilon_+ - \epsilon_-) \phi_{\mathbf{k}} |\mathbf{k}, \mathbf{p}\rangle + \sum_{|\mathbf{k}|, |\mathbf{k}'| > k_F} |\mathbf{k}, \mathbf{p}\rangle \langle \mathbf{k}, \mathbf{p} | V | \mathbf{k}', \mathbf{p} \rangle \phi_{\mathbf{k}'} \quad (10.102)$$

where $\epsilon_{\pm} = \epsilon_{\mathbf{k}\pm\mathbf{p}/2}$. Now assume that $\langle \mathbf{k}, \mathbf{p} | V | \mathbf{k}', \mathbf{p} \rangle \phi_{\mathbf{k}'} = -g_0/\mathcal{U}$, then

$$E_{\mathbf{p}} \phi_{\mathbf{k}} = (\epsilon_+ - \epsilon_-) \phi_{\mathbf{k}} - \frac{g_0}{\mathcal{U}} \sum_{0 < \epsilon_{\pm} < \omega_D} \phi_{\mathbf{k}'} \quad (10.103)$$

where ω_D is the cutoff energy above which the phonon mediation of the electron-electron interaction no longer makes the effective interaction attractive. More accurately, we say that Cooper pairs can only form within the interval $-\omega_D \leq \epsilon \leq \omega_D$, so both above and below the Fermi-energy. We can rearrange and obtain

$$1 - \frac{g_0}{U} \sum_{0 < \epsilon_{\pm} < \omega_D} \frac{1}{\epsilon_+ - \epsilon_- - E_{\mathbf{p}}} = 0 \quad (10.104)$$

It is now convenient to define a function $G^{-1}(E_{\mathbf{p}}, \mathbf{p})$ whose zeros we would like to find:

$$G^{-1}(z, \mathbf{p}) = 1 - g_0 \chi_0(z, \mathbf{p}), \quad \chi_0(z, \mathbf{p}) = \frac{1}{U} \sum_{0 < \epsilon_{\pm} < \omega_D} \frac{1}{\epsilon_+ - \epsilon_- - z} \quad (10.105)$$

for a quadratic dispersion we can solve this to obtain

$$G^{-1}(E, p) = 1 - \frac{g_0 N(0)}{2} \ln \left(\frac{2\omega_D}{v_F p - E} \right) \quad (10.106)$$

thus we satisfy the Schrödinger equation by setting this equal to zero:

$$E_{\mathbf{p}} = -2\omega_D \exp \left[-\frac{2}{g_0 N(0)} \right] + v_F p \quad (10.15)$$

the linear spectrum is a signature of a collective bosonic mode. Notice how this energy is negative, which implies that it is energetically *beneficial* to form Cooper Pairs – hence the name “Cooper Instability”.

10.3.3 The BCS Hamiltonian

However, this Cooper-pair state that we discussed in the previous section is not sufficient to understand superconductivity. We need to create a coherent state with the Cooper pair operator:

$$|\psi_{\text{BCS}}\rangle = \exp(\Lambda^\dagger) |0\rangle = \prod_{\mathbf{k}} (1 + \phi_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}) |0\rangle \quad (10.16)$$

where we see that the exponential simplifies because we are looking at fermionic raising and lowering operators. Notice how we have let $\mathbf{p} = 0$ for simplicity – it seems this still describes the physics quite well. Notice how this state has the peculiar property that

$$\langle \psi_{\text{BCS}} | c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} | \psi_{\text{BCS}} \rangle \sim \phi_{\mathbf{k}} \neq 0 \quad (10.17)$$

The question now arises: what kind of a Hamiltonian describes a system where $|\psi_{\text{BCS}}\rangle$ is an eigenstate. The answer is the BCS Hamiltonian:

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k},\mathbf{k}'} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \quad (10.18)$$

$V_{\mathbf{k},\mathbf{k}'}$ can take on a whole variety of different forms, however commonly one uses

$$V_{\mathbf{k},\mathbf{k}'} = \begin{cases} -\frac{g_0}{U} & |\epsilon_{\mathbf{k}}| < \omega_D \\ 0 & \text{otherwise} \end{cases} \quad (10.19)$$

This is the *s-wave manifestation*. We can write this in a form that it is ready for what is to come – a Hubbard-Stratonovich transformation:

$$H = \sum_{|\epsilon_{\mathbf{k}}| < \omega_D, \sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} - \frac{g_0}{V} A^\dagger A \quad (10.20)$$

where $A = \sum_{|\epsilon_{\mathbf{k}}| < \omega_D} c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow}$.

Sometimes papers refer to s-wave superconductors and this is what is meant.

10.3.4 Mean-Field Description of the Condensate

As we saw, the BCS-ground state exhibits so-called *off-diagonal-long-range order*, which is the fact that Equation 10.17 is non-zero. The pair operator A is extensive, and in the superconducting state $\langle A \rangle \propto \mathcal{U}$, however, the pair *density*

$$\Delta = |\Delta|e^{i\phi} = -\frac{g_0}{\mathcal{U}} \langle A \rangle \quad (10.21)$$

is intensive. This value, Δ , is the *order parametre* of superconductivity and contrary to other order parametres it has both a modulus and a phase. The fact that it is an order parametre is quite intuitive: above the critical temperature (*i.e* when are we not in the superconducting state) there is no off-diagonal-long-range order, whereas below the critical temperature there is. It turns out that $\Delta \sim \sqrt{T_C - T}$ close to T_C as one may expect from an order parametre.

If we express $A = \langle A \rangle + \delta A = \Delta + \delta A$ then we obtain that

$$-\frac{g_0}{\mathcal{U}} A^\dagger A = \bar{\Delta} A + A^\dagger \Delta + \mathcal{U} \frac{\bar{\Delta} \Delta}{g_0} + \frac{g_0}{\mathcal{U}} \delta A^\dagger \delta A \quad (10.22)$$

it turns out that the last term can be ignored in the thermodynamic limit, as it is $O(1)$ whereas the other terms are $O(\mathcal{U})$, which means that in the thermodynamic limit we can rewrite the BCS Hamiltonian as a mean-field Hamiltonian and not omit any important physics:

$$H_{\text{MFT}} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \sum_{\mathbf{k}} \left[\bar{\Delta} c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} + c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger \Delta \right] + \frac{\mathcal{U}}{g_0} \bar{\Delta} \Delta \quad (10.23)$$

where Δ is to be determined self-consistently by minimising the free energy.

10.3.5 Physical Pictures of BCS Theory

Pairs as Spins

The pairing term in the mean-field Hamiltonian:

$$H_P(\mathbf{k}) = \bar{\Delta} c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} + c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger \Delta \quad (10.24)$$

can on the one hand be thought of as the following

$$e^- + e^- \rightleftharpoons \text{pair}^{2-} \quad (10.25)$$

however, it can also be beneficial to instead think of it as the creation of a cooper pair and a hole (we move one of the electrons to the right-hand-side of the process above), so we instead think of $c_{-\mathbf{k}\downarrow} = h_{\mathbf{k}}^\dagger$, which means we write the pairing term as

$$H_P(\mathbf{k}) = \bar{\Delta} h_{\mathbf{k}}^\dagger c_{\mathbf{k}\uparrow} + c_{\mathbf{k}\uparrow}^\dagger h_{\mathbf{k}} \Delta \quad (10.26)$$

where now

$$e^- \rightleftharpoons \text{pair}^{2-} + h^+ \quad (10.27)$$

This process is referred to as *Andreev Reflection*. Following this we should associate the particle energy with $\epsilon_{\mathbf{k}}$ but the hole energy with $-\epsilon_{-\mathbf{k}}$. The dispersions intersect precisely at the Fermi-surface, which means that the mixing of these states (which occurs due to the Andreev Reflection) will open a gap, eliminating the Fermi-Surface. That is, in the particle-hole basis

$$\psi_{\mathbf{k}} = \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ h_{\mathbf{k}\downarrow} \end{pmatrix} \quad (10.28)$$

the Hamiltonian reads

$$H = \begin{pmatrix} \epsilon_{\mathbf{k}} & \Delta \\ \bar{\Delta} & -\epsilon_{\mathbf{k}} \end{pmatrix} \rightsquigarrow E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2 + |\Delta|^2} \quad (10.29)$$

Nambu Spinors

The basis I referred to as the particle-hole basis is actually called a *Nambu Spinor*:

$$\psi_{\mathbf{k}} = \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix}, \quad \psi_{\mathbf{k}}^\dagger = \begin{pmatrix} c_{\mathbf{k}\uparrow}^\dagger & c_{-\mathbf{k}\downarrow} \end{pmatrix} \quad (10.30)$$

which allows us to write the full Hamiltonian as

$$H = \sum_{\mathbf{k}} \psi_{\mathbf{k}}^\dagger (\mathbf{h}_{\mathbf{k}} \cdot \boldsymbol{\tau}) \psi_{\mathbf{k}} + U \frac{\bar{\Delta}\Delta}{g_0} \quad (10.31)$$

where

$$\mathbf{h}_{\mathbf{k}} = \begin{pmatrix} \text{Re}\Delta \\ -\text{Im}\Delta \\ \epsilon_{\mathbf{k}} \end{pmatrix} = \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \epsilon_{\mathbf{k}} \end{pmatrix} \quad (10.32)$$

and $\boldsymbol{\tau}$ is the vector of Pauli-Matrices. Thus we see that $\mathbf{h}_{\mathbf{k}}$ behaves like a Zeeman field, but in isospin space.

Anderson's Domain-Wall Interpretation of BCS-Theory

Let us define an isospin operator $\boldsymbol{\tau}_{\mathbf{k}} \equiv \psi_{\mathbf{k}}^\dagger \boldsymbol{\tau} \psi_{\mathbf{k}}$, which has the properties of a spin- $\frac{1}{2}$ operator acting in charge space. The z-component can be written out as

$$\tau_{3\mathbf{k}} = n_{\mathbf{k}\uparrow} + n_{-\mathbf{k}\downarrow} - 1 \quad (10.33)$$

Thus we can define its eigenvectors as up- and down-spin states in isospin space:

$$\begin{aligned} \tau_{3\mathbf{k}} |\uparrow_{\mathbf{k}}\rangle &= |\uparrow_{\mathbf{k}}\rangle, & |\uparrow_{\mathbf{k}}\rangle &= |2\rangle = c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger |0\rangle \\ \tau_{3\mathbf{k}} |\downarrow_{\mathbf{k}}\rangle &= -|\downarrow_{\mathbf{k}}\rangle, & |\downarrow_{\mathbf{k}}\rangle &= |0\rangle \end{aligned} \quad (10.34)$$

Notice how in a normal metal $\Delta = 0$, so only τ_3 appears in Equation 10.31, so we are in the up-state below the Fermi-surface, but in the down-state above the Fermi-surface. In the superconducting phase the transition from fully up to fully down will be softened, as is shown below

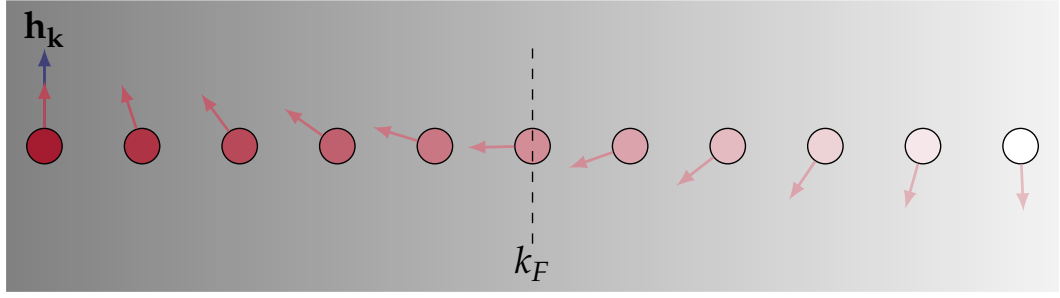


Figure 10.2: The isospin along the z -axis is represented as the vertical component of the arrow. Inside the Fermi surface the up-state is occupied, outside the Fermi surface it is the down-state that is occupied

In a normal metal a quasiparticle arises when we flip the isospin of a state, for instance if we are outside the Fermi surface we can make a quasiparticle by exciting a particular state into the up-state. In superconductors we define a *Bogoliubov Quasiparticle* as the state where we have flipped the isospin of a particle, however, now there are also x and y components that are flipped.

As we saw, the quasiparticle energy is $E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2 + |\Delta|^2}$, which allows us to define a unit vector in isospin space

$$\hat{n}_{\mathbf{k}} = \frac{1}{E_{\mathbf{k}}} \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \epsilon_{\mathbf{k}} \end{pmatrix} \quad (10.35)$$

The $U(1)$ symmetry of our system allows us to freely choose a phase of Δ , so we can effectively eliminate Δ_2 . Thus the arrows in Figure 10.2 are the nonzero components of $\hat{n}_{\mathbf{k}}$. We can define

$$\cos \theta_{\mathbf{k}} \equiv \frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} = \langle \tau_{3\mathbf{k}} \rangle, \quad \sin \theta_{\mathbf{k}} \equiv \frac{\Delta}{E_{\mathbf{k}}} = \langle \tau_{1\mathbf{k}} \rangle \quad (10.36)$$

Notice how

$$\sin \theta_{\mathbf{k}} = \langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger + c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \rangle = 2 \langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger \rangle \quad (10.37)$$

because we have chosen $\langle \tau_{2\mathbf{k}} \rangle = \langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger - c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \rangle = 0$ and hence Equation 10.21 becomes

$$\Delta = g_0 \int_{|\epsilon_{\mathbf{k}}| < \omega_D} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\Delta}{2\sqrt{\epsilon_{\mathbf{k}}^2 + \Delta^2}} \quad (10.38)$$

which is the *BCS Gap Equation* at zero temperature. If we assume that the density of states is constant around the Fermi-level (very common assumption for superconductivity) we get

$$1 = g_0 N(0) \int_{-\omega_D}^{\omega_D} d\epsilon \frac{1}{2\sqrt{\epsilon^2 + \Delta^2}} = g_0 N(0) \sinh^{-1} \left(\frac{\omega_D}{\Delta} \right) \approx g_0 N(0) \ln \left(\frac{2\omega_D}{\Delta} \right) \quad (10.39)$$

so, in the superconducting ground state, the BCS gap is given by

$$\Delta = 2\omega_D \exp \left(-\frac{1}{g_0 N(0)} \right) \quad (10.40)$$

notice the disappearance of the factor 2, which we saw in Equation 10.15.

10.3.6 The BCS Ground State

We would now like to construct a state whose isospin is always aligned with the Weiss-field ($\mathbf{h}_\mathbf{k}$). This is done by simply taking the linear combination

This is a normalised version of the coherent state discussed previously

$$|\text{BCS}\rangle = \prod_{\mathbf{k}} \left(\cos \frac{\theta_{\mathbf{k}}}{2} + \sin \frac{\theta_{\mathbf{k}}}{2} c_{-\mathbf{k}\downarrow}^\dagger c_{\mathbf{k}\uparrow}^\dagger \right) |0\rangle \quad (10.41)$$

the coefficients are commonly written as

$$|\text{BCS}\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{-\mathbf{k}\downarrow}^\dagger c_{\mathbf{k}\uparrow}^\dagger) |0\rangle \quad (10.42)$$

where

$$u_{\mathbf{k}} = \sqrt{\frac{1}{2} \left[1 + \frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} \right]}, \quad v_{\mathbf{k}} = \sqrt{\frac{1}{2} \left[1 - \frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} \right]} \quad (10.43)$$

Notice how the BCS-ground state breaks the $U(1)$ -symmetry which the BCS-Hamiltonian possesses, let $c_{\mathbf{k}\sigma}^\dagger \rightarrow e^{i\alpha} c_{\mathbf{k}\sigma}^\dagger$, then (ignoring the normalisation):

$$|\text{BCS}\rangle \rightarrow |\alpha\rangle = \prod_{\mathbf{k}} (1 + e^{2i\alpha} \phi_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger) |0\rangle = \sum \frac{e^{i2n\alpha}}{\sqrt{n!}} |n\rangle \quad (10.44)$$

which means that $\Delta = -\frac{g_0}{U} \sum_{\mathbf{k}} \langle \alpha | c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} | \alpha \rangle$ obtains a phase:

$$\Delta \rightarrow e^{2i\alpha} \Delta \quad (10.45)$$

Additionally we have that

$$\hat{N}|\alpha\rangle = \sum \frac{1}{\sqrt{n!}} 2ne^{i2n\alpha} |\alpha\rangle = -i \frac{d}{d\alpha} |\alpha\rangle \quad (10.46)$$

So that $\hat{N} \equiv -i \frac{d}{d\alpha}$. This is exactly like the relationship between the momentum and position operators.

10.3.7 Quasiparticle Excitations in BCS Theory

Let us diagonalise the BCS-Hamiltonian

$$H = \sum_{\mathbf{k}} \psi_{\mathbf{k}}^\dagger (\mathbf{h}_{\mathbf{k}} \cdot \boldsymbol{\tau}) \psi_{\mathbf{k}} + U \frac{\bar{\Delta} \Delta}{g_0} \quad (10.47)$$

We are interested in diagonalising the first term (Δ is just a number), which we can do by writing this term using the previously defined vector $\hat{n}_{\mathbf{k}}$:

$$\mathbf{h}_{\mathbf{k}} \cdot \boldsymbol{\tau} = E_{\mathbf{k}} \hat{n}_{\mathbf{k}} \cdot \boldsymbol{\tau} \quad (10.48)$$

so we are interested in the eigenvectors of $\hat{n}_{\mathbf{k}} \cdot \boldsymbol{\tau}$:

$$\hat{n}_{\mathbf{k}} \cdot \boldsymbol{\tau} \begin{pmatrix} u_{\mathbf{k}} \\ v_{\mathbf{k}} \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}} \\ v_{\mathbf{k}} \end{pmatrix}, \quad \hat{n}_{\mathbf{k}} \cdot \boldsymbol{\tau} \begin{pmatrix} -v_{\mathbf{k}}^* \\ u_{\mathbf{k}}^* \end{pmatrix} = - \begin{pmatrix} -v_{\mathbf{k}}^* \\ u_{\mathbf{k}}^* \end{pmatrix} \quad (10.49)$$

thus by defining the unitary matrix

$$U_{\mathbf{k}} = \begin{pmatrix} u_{\mathbf{k}} & -v_{\mathbf{k}}^* \\ v_{\mathbf{k}} & u_{\mathbf{k}}^* \end{pmatrix} \quad (10.50)$$

we can diagonalise the Hamiltonian

$$U_{\mathbf{k}}^\dagger (\hat{n}_{\mathbf{k}} \cdot \boldsymbol{\tau}) U_{\mathbf{k}} = \tau_3 \quad (10.51)$$

This transformation is a *Bogoliubov Transformation*, which we encountered in CMT1:

$$\begin{pmatrix} a_{\mathbf{k}\uparrow} \\ a_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}}^* & v_{\mathbf{k}}^* \\ -v_{\mathbf{k}} & u_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix} \quad (10.52)$$

In this basis the Hamiltonian reads

$$H = \sum_{\mathbf{k}} E_{\mathbf{k}} (a_{\mathbf{k}\uparrow}^\dagger a_{\mathbf{k}\uparrow} - a_{-\mathbf{k}\downarrow} a_{-\mathbf{k}\downarrow}^\dagger) + \mathcal{U} \frac{\bar{\Delta} \Delta}{g_0} \quad (10.53)$$

which we can write in a more familiar form

$$H = \sum_{\mathbf{k}\sigma} E_{\mathbf{k}} \left(a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} - \frac{1}{2} \right) + \mathcal{U} \frac{\bar{\Delta} \Delta}{g_0} \quad (10.54)$$

We can obtain the density of Bogoliubov quasiparticles per spin through the relation $N_S(E) dE = N_N(0) d|\epsilon|$, where $N_N(0) = 2N(0)$:

$$N_S(E_{\mathbf{k}}) = N_N(0) \frac{d|\epsilon_{\mathbf{k}}|}{dE_{\mathbf{k}}} = N_n(0) \left(\frac{E_{\mathbf{k}}}{\sqrt{E_{\mathbf{k}}^2 - |\Delta|^2}} \right) \theta(E - |\Delta|) \quad (10.55)$$

which is plotted below

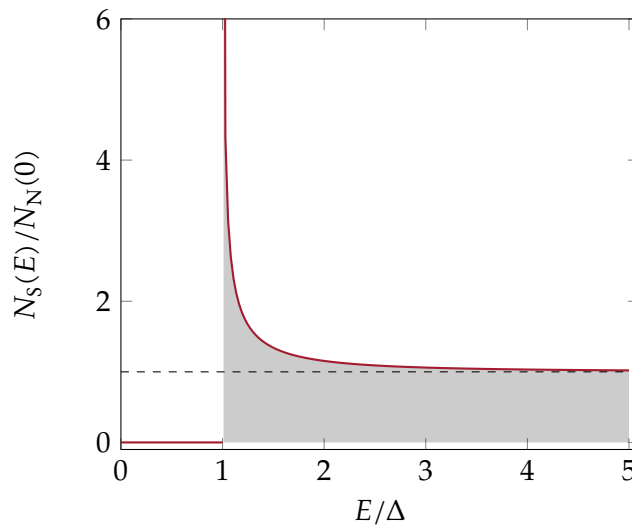


Figure 10.3: BCS Theory's prediction of the density of states of a superconductor with a square root discontinuity at $E = \Delta$: the coherence peak

IO.3.8 Path Integral Formulation

As anticipated the BCS-path integral looks like

$$Z = \int \mathcal{D}[\bar{c}, c] e^{-S}, \quad S = \int_0^\beta d\tau \sum_{\mathbf{k}\sigma} \bar{c}_{\mathbf{k}\sigma} (\partial_\tau + \epsilon_{\mathbf{k}}) c_{\mathbf{k}\sigma} - \frac{g_0}{\mathcal{V}} \bar{A} A \quad (10.56)$$

When dealing with superconductivity we must always keep in mind that $|\epsilon_{\mathbf{k}}| < \omega_D$. The Hubbard-Stratonovich transformation removes the quartic term:

$$-g \bar{A} A \rightarrow \bar{\Delta} A + \bar{A} \Delta + \frac{\mathcal{V}}{g_0} \bar{\Delta} \Delta \quad (10.57)$$

where $\Delta(\tau)$ are fluctuating complex fields:

$$Z = \int \mathcal{D}[\Delta, c] e^{-S}, \quad S = \int_0^\beta d\tau \sum_{\mathbf{k}\sigma} \bar{c}_{\mathbf{k}\sigma} (\partial_\tau + \epsilon_{\mathbf{k}}) c_{\mathbf{k}\sigma} + \bar{\Delta} A + \bar{A} \Delta + \frac{\mathcal{V}}{g_0} \bar{\Delta} \Delta \quad (10.58)$$

As expected this is expressed nicely with Nambu-Spinors

$$Z = \int \mathcal{D}[\Delta, \psi] e^{-S}, \quad S = \int_0^\beta d\tau \sum_{\mathbf{k}\sigma} \bar{\psi}_{\mathbf{k}} (\partial_\tau + \underline{h}_{\mathbf{k}}) \psi_{\mathbf{k}} + \frac{\mathcal{V}}{g_0} \bar{\Delta} \Delta \quad (10.59)$$

where $\underline{h}_{\mathbf{k}} = \epsilon_{\mathbf{k}} \tau_3 + \Delta_1 \tau_1 + \Delta_2 \tau_2$ and $\Delta = \Delta_1 - i\Delta_2$. We can integrate out the fermionic fields:

$$Z = \int \mathcal{D}[\bar{\Delta}, \Delta] e^{-S_E[\bar{\Delta}, \Delta]}, \quad e^{-S_E[\bar{\Delta}, \Delta]} = \prod_{\mathbf{k}} \det[\partial_\tau + \underline{h}_{\mathbf{k}}(\tau)] e^{-\mathcal{V} \int_0^\beta d\tau \frac{\bar{\Delta} \Delta}{g_0}} \quad (10.60)$$

and so

$$S_E[\bar{\Delta}, \Delta] = \mathcal{V} \int_0^\beta d\tau \frac{\bar{\Delta} \Delta}{g_0} - \sum_{\mathbf{k}} \text{tr} \ln[\partial_\tau + \underline{h}_{\mathbf{k}}] \quad (10.61)$$

As we discussed previously, for superconductors in the thermodynamic limit we can effectively replace the Hamiltonian with its mean-field equivalent, because the δA terms disappeared for sufficiently large volumes. Thus we can treat Δ as its expectation value, which we will assume is independent of τ . Thus the mean field partition function $Z \approx Z_{\text{BCS}} = e^{-S_E[\bar{\Delta}_0, \Delta_0]}$ is

$$Z_{\text{BCS}} = \prod_{\mathbf{k}} \det[\partial_\tau + \underline{h}_{\mathbf{k}}] \exp\left(-\frac{\mathcal{V}\beta}{g_0} \bar{\Delta} \Delta\right) \quad (10.62)$$

where we have omitted the 0-subscript on Δ for brevity. As usual it is easier if we work in Matsubara-Fourier space

$$Z_{\text{BCS}} = \int \prod_{\mathbf{k}n} d\bar{\psi}_{\mathbf{k}n} d\psi_{\mathbf{k}n} e^{-S_{\text{MFT}}[\bar{\psi}_{\mathbf{k}n}, \psi_{\mathbf{k}n}]}, \quad (10.63)$$

$$S_{\text{MFT}}[\bar{\psi}_{\mathbf{k}n}, \psi_{\mathbf{k}n}] = \sum_{\mathbf{k}n} \bar{\psi}_{\mathbf{k}n} (-i\omega_n + \underline{h}_{\mathbf{k}}) \psi_{\mathbf{k}n} + \beta \mathcal{V} \frac{\bar{\Delta} \Delta}{g_0}$$

where $\psi_{\mathbf{k}}(\tau) = \frac{1}{\sqrt{\beta}} \sum_n \psi_{\mathbf{k}n} e^{-i\omega_n \tau}$. We can integrate out the fermionic fields to obtain

$$Z_{\text{BCS}} = \prod_{\mathbf{k}n} [\omega_n^2 + \epsilon_{\mathbf{k}}^2 + |\Delta|^2] e^{-\frac{\beta \mathcal{V} |\Delta|^2}{g_0}} = e^{-S_E} \quad (10.64)$$

and hence we have an expression for the free energy

$$F[\Delta, T] = -T \sum_{\mathbf{k}n} \ln [\omega_n^2 + \epsilon_{\mathbf{k}}^2 + |\Delta|^2] + \mathcal{V} \frac{|\Delta|^2}{g_0} \quad (10.65)$$

BCS Gap Equations: Revisited

The free energy is minimised when $\frac{\partial F}{\partial \Delta} = 0$, which gives us

$$\frac{1}{g_0} = \frac{1}{\beta \mathcal{V}} \sum_{\mathbf{k}n} \frac{1}{\omega_n^2 + E_{\mathbf{k}}^2} \quad (10.66)$$

which is the first BCS Gap Equation, but we can do better than this, by performing the Matsubara sum.

$$\frac{1}{\beta} \sum_n \frac{1}{\omega_n^2 + E_{\mathbf{k}}^2} \xrightarrow[\text{Integral}]{\text{Contour}} \frac{\tanh\left(\frac{\beta E_{\mathbf{k}}}{2}\right)}{2E_{\mathbf{k}}} \quad (10.67)$$

and so we obtain the second BCS gap equation

$$\frac{1}{g_0} = \int_{|\epsilon_{\mathbf{k}}| < \omega_D} \frac{d^3 \mathbf{k}}{(2\pi)^3} \left[\frac{\tanh(\beta E_{\mathbf{k}}/2)}{2E_{\mathbf{k}}} \right] \quad (10.68)$$

10.3.9 Nambu-Gor'kov Green's Function

Let us define a Green's function (matrix) which we make by taking the time-ordered expectation value of two Nambu-spinors:

$$\underline{G}_{\alpha\beta}(\mathbf{k}, \tau) = -\left\langle \mathbb{T} \psi_{\mathbf{k}\alpha}(\tau) \psi_{\mathbf{k}\beta}^\dagger(0) \right\rangle \quad (10.69)$$

One often calls the off-diagonal components F :

$$F(\mathbf{k}, \tau) = -\left\langle \mathbb{T} c_{\mathbf{k}\uparrow}(\tau) c_{-\mathbf{k}\downarrow}(0) \right\rangle, \quad \bar{F}(\mathbf{k}, \tau) = -\left\langle \mathbb{T} c_{-\mathbf{k}\downarrow}^\dagger(\tau) c_{\mathbf{k}\uparrow}^\dagger(0) \right\rangle \quad (10.70)$$

Notice how these are non-zero for the superconducting state, due to the off-diagonal-long-range order. We can write this out in Matsubara space

$$\underline{G}(k) = \frac{1}{(i\omega_n)^2 - E_{\mathbf{k}}^2} (i\omega_n + \Delta_1 \tau_1 + \Delta_2 \tau_2 + \epsilon_{\mathbf{k}} \tau_3) \quad (10.71)$$

This can be obtained by our previous methods or by using the Dyson Equation where the self-energy is given by the sum over even numbers of Andreev Reflections.

Note that the diagonal terms are

$$\text{diag } G_k = \frac{1}{(i\omega_n)^2 - E_{\mathbf{k}}^2} \begin{pmatrix} i\omega_n + \epsilon_{\mathbf{k}} \\ i\omega_n - \epsilon_{\mathbf{k}} \end{pmatrix} \quad (10.72)$$

Notice how

$$\frac{1}{2} \left(\frac{1}{i\omega_n + E_{\mathbf{k}}} + \frac{1}{i\omega_n - E_{\mathbf{k}}} \right) = \frac{i\omega_n}{(i\omega_n)^2 - E_{\mathbf{k}}^2} \quad (10.73)$$

$$-\frac{\epsilon_{\mathbf{k}}}{2E_{\mathbf{k}}} \left(\frac{1}{i\omega_n + E_{\mathbf{k}}} - \frac{1}{i\omega_n - E_{\mathbf{k}}} \right) = \frac{\epsilon_{\mathbf{k}}}{(i\omega_n)^2 - E_{\mathbf{k}}^2} \quad (10.74)$$

which means that

$$\frac{i\omega_n + \epsilon_{\mathbf{k}}}{(i\omega_n)^2 - E_{\mathbf{k}}^2} = \frac{v_{\mathbf{k}}^2}{i\omega_n + E_{\mathbf{k}}} + \frac{u_{\mathbf{k}}^2}{i\omega_n - E_{\mathbf{k}}} \quad (10.75)$$

$$\frac{i\omega_n - \epsilon_{\mathbf{k}}}{(i\omega_n)^2 - E_{\mathbf{k}}^2} = \frac{u_{\mathbf{k}}^2}{i\omega_n + E_{\mathbf{k}}} + \frac{v_{\mathbf{k}}^2}{i\omega_n - E_{\mathbf{k}}} \quad (10.76)$$

which tells us that

$$\text{diag } \mathcal{G}_k = \begin{pmatrix} u_{\mathbf{k}}^2 & v_{\mathbf{k}}^2 \\ v_{\mathbf{k}}^2 & u_{\mathbf{k}}^2 \end{pmatrix} \begin{pmatrix} (i\omega - E_{\mathbf{k}})^{-1} \\ (i\omega + E_{\mathbf{k}})^{-1} \end{pmatrix} \quad (10.77)$$

Tunneling Density of States and Coherence Factors

Let us consider the particle part of the Nambu-Gor'kov Green's function, so \mathcal{G}_{11} . If we analytically continue it and take the imaginary part we get

$$A(\mathbf{k}, \omega) = \frac{1}{\pi} \text{Im } \mathcal{G}_{11}(\mathbf{k}, \omega - i\delta) = u_{\mathbf{k}}^2 \delta(\omega - E_{\mathbf{k}}) + v_{\mathbf{k}}^2 \delta(\omega + E_{\mathbf{k}}) \quad (10.78)$$

so we see that $u_{\mathbf{k}}^2$ describes the probability to create a positive energy quasiparticle, whereas $v_{\mathbf{k}}^2$ describes the probability to create a negative energy state.

IO.3.10 From BCS To Ginzburg-Landau

To derive Ginzburg-Landau theory, let us rewrite Equation 10.61 in a way that we can easily expand in orders of the order-parametre:

$$S[\bar{\Delta}, \Delta] = \mathcal{U} \int_0^\beta d\tau \frac{\bar{\Delta} \Delta}{g_0} - \text{tr} \ln [-\underline{G}_0^{-1} (1 + \underline{G}_0 \underline{\Delta})] \quad (10.79)$$

where

$$-\underline{G}_0^{-1} = -i\omega_n + \epsilon_{\mathbf{k}} \tau_3, \quad \underline{\Delta} = \Delta_1 \tau_1 + \Delta_2 \tau_2 \quad (10.80)$$

now, assuming that Δ is a small parametre, then we can expand the logarithm as

$$-\text{tr} \ln [1 + \underline{G}_0 \underline{\Delta}] \approx -\text{tr} [\underline{G}_0 \underline{\Delta}] + \frac{1}{2} \text{tr} [\underline{G}_0 \underline{\Delta} \underline{G}_0 \underline{\Delta}] + O(\Delta^4) \quad (10.81)$$

let's ignore the fourth order term and calculate the second order term. It turns out that

$$\frac{1}{2} \text{tr} [\underline{G}_0 \underline{\Delta} \underline{G}_0 \underline{\Delta}] = \frac{1}{\beta \mathcal{U}} \sum_{kk'} G_{0\uparrow}(k) \Delta_q G_{0\downarrow}(k') \bar{\Delta}_{-q} \quad (10.82)$$

so up to fourth order in Δ the effective action is

$$S[\bar{\Delta}, \Delta] = \sum_q \left(\frac{1}{g_0} - \frac{1}{\beta V} \sum_k G_0(k) G_0(q - k) \right) |\Delta_q|^2 + \frac{1}{4} |\Delta|^4 \quad (10.83)$$

where u comes from $\text{tr} [\underline{G} \underline{\Delta} \underline{G} \underline{\Delta}]$ terms. Thus by assuming that Δ is small we have obtained Ginzburg-Landau theory, which tells us a lot about the physics of superconductivity, however, instead of it being a lucky guess, we have derived it from a microscopic theory.

10.3.11 *Anderson-Higgs Mechanism in BCS-Theory*

The full Nambu-Gor'kov Green's function is

$$\tilde{\underline{G}}^{-1} = -\tau_0 \partial_\tau - \tau_1 \Delta_0 - \tau_3 \left(e\tilde{\phi} + \frac{1}{2m} (-i\nabla - e\tau_3 \tilde{\mathbf{A}})^2 - \mu \right) \quad (10.84)$$

where we have gauged Δ 's phase out by means of a unitary transformation, which leaves the action unchanged. The transformation however let

$$\Delta \rightarrow \Delta_0 = |\Delta|, \quad \mathbf{A} \rightarrow \mathbf{A} - e^{-1} \nabla \theta \equiv \tilde{\mathbf{A}}, \quad \phi \rightarrow \phi + ie^{-1} \partial_\tau \theta \equiv \tilde{\phi} \quad (10.85)$$

Let us split this up into three parts depending on which order of \tilde{A}_μ appears:

$$\tilde{\underline{G}}^{-1} = \tilde{\underline{G}}_0^{-1} - \underline{\chi}_1 - \underline{\chi}_2 \quad (10.86)$$

where

$$\begin{aligned} \tilde{\underline{G}}_0^{-1} &\equiv -\tau_0 \partial_\tau - \tau_3 \left(\frac{(-i\nabla)^2}{2m} - \mu \right) - \tau_1 \Delta_0 \\ \underline{\chi}_1 &\equiv \tau_3 e\tilde{\phi} - \frac{e}{2m} \tau_0 [-i\nabla, \tilde{\mathbf{A}}]_+ \\ \underline{\chi}_2 &\equiv \tau_3 \frac{e^2 \tilde{A}^2}{2m} \end{aligned} \quad (10.87)$$

the action can be expanded using $\ln(1-x) \approx -x - \frac{1}{2}x^2$:

$$S = -\text{tr} \ln[-\tilde{\underline{G}}^{-1}] \approx -\text{tr} \ln[-\tilde{\underline{G}}_0^{-1}] + \text{tr} \ln[\tilde{\underline{G}}_0 \underline{\chi}_1] + \text{tr} \ln[\tilde{\underline{G}}_0 \underline{\chi}_2] + \frac{1}{2} \text{tr} \ln[(\tilde{\underline{G}}_0 \underline{\chi}_1)^2] \quad (10.88)$$

First Order in the Electromagnetic-Field

The first order term

$$S^{(1)} = \frac{1}{\beta \mathcal{V}} \sum_k \text{tr} [\tilde{\underline{G}}_0(k) \underline{\chi}_1(k, k)] = \frac{e}{\beta \mathcal{V}} \sum_k \text{tr} [\tilde{\underline{G}}_0(k) (\tau_3 \phi_0 - m^{-1} \tau_0 \mathbf{k} \cdot \mathbf{A}_0)] \quad (10.89)$$

where we have used that $A_\mu = \tilde{A}_\mu$ at $q = 0$. $\tilde{\underline{G}}_0^{-1}$ is even in \mathbf{k} , which implies that the $\mathbf{k} \cdot \mathbf{A}_0$ term integrates to zero leaving only

$$\begin{aligned} S^{(1)} &= \frac{e\phi_0}{\beta \mathcal{V}} \sum_k \text{tr} [\tilde{\underline{G}}_0(k) \tau_3] \\ &= \frac{e\phi_0}{\mathcal{V}} \sum_{\mathbf{k}} \frac{1}{\beta} \sum_{i\omega_n} \left(\frac{i\omega_n + \epsilon_k}{(i\omega_n)^2 - E_k^2} - \frac{i\omega_n - \epsilon_k}{(i\omega_n)^2 - E_k^2} \right) \end{aligned} \quad (10.90)$$

performing the Matsubara sum gives us $S^{(1)} \approx en \int d\tau d\mathbf{r} \phi(\mathbf{r}, \tau)$, which is cancelled by the corresponding term from the positive ionic background – we are assuming that the superconductor is neutral.

Second Order in the Electromagnetic-Field

There are two second order terms, the first of which is quite simple as the calculation is exactly what we did, but replacing $\phi_0 \rightsquigarrow \frac{eA_0}{2m}$:

$$\text{tr} [\tilde{\underline{G}}_0 \underline{\chi}_2] = \frac{ne^2}{2m} A_0^2 \quad (10.91)$$

That is the *diamagnetic term*. Now for the paramagnetic term

$$\begin{aligned} \frac{1}{2} \text{tr} [\tilde{\underline{G}}_0 \underline{\chi}_1 \tilde{\underline{G}}_0 \underline{\chi}_1] &= \frac{1}{2\beta} \sum_{kq} \text{tr} [\tilde{\underline{G}}_0(k) \tau_3 (e\tilde{\phi}_q) \tilde{\underline{G}}_0(k) \tau_3 (e\tilde{\phi}_{-q})] + \\ &\quad \frac{1}{m^2} \text{tr} [\tilde{\underline{G}}_0(k) \tilde{\underline{G}}_0(k) (\mathbf{k} \cdot \tilde{\mathbf{A}}_q) (\mathbf{k} \cdot \tilde{\mathbf{A}}_{-q})] \end{aligned} \quad (10.92)$$

where we've ignore the q in $\underline{G}_0(k+q)$, because this just gives us an extra term that doesn't change the physics of the problem. This evaluates to

$$\int d\tau d\mathbf{r} [-c_1 e^2 \tilde{\phi}^2(\mathbf{r}, \tau) + c_2 e^2 \tilde{\mathbf{A}}^2(\mathbf{r}, \tau)] = \int d\tau d\mathbf{r} [-c_1 (\partial_\tau \theta - ie\phi)^2 + c_2 (\nabla \theta - e\mathbf{A})^2] \quad (10.93)$$

where $c_1 \equiv \nu_F$ and $c_2 \equiv \frac{n_s}{2m}$ and n_s is the superfluid density.

Now we assume that we are at temperatures high enough that we ignore the $\partial_\tau \theta$ term, and let us reintroduce the $(\nabla \times \mathbf{A})$ term which we in reality should have included from the beginning, then the action reads

$$\begin{aligned} S[\mathbf{A}, \theta] &= \frac{\beta}{2} \int d\mathbf{r} \left[\frac{n_s}{m} (\nabla \theta - e\mathbf{A})^2 + \frac{1}{\mu_0} (\nabla \times \mathbf{A})^2 \right] \\ &= \frac{\beta}{2} \sum_{\mathbf{q}} \left[\frac{n_s}{m} (i\mathbf{q}\theta_{\mathbf{q}} - e\mathbf{A}_{\mathbf{q}}) \cdot (-i\mathbf{q}\theta_{-\mathbf{q}} - e\mathbf{A}_{-\mathbf{q}}) \right] + \frac{1}{\mu_0} (\mathbf{q} \times \mathbf{A}_{\mathbf{q}}) (\mathbf{q} \times \mathbf{A}_{-\mathbf{q}}) \end{aligned} \quad (10.94)$$

Integrating out θ we get

$$S[\mathbf{A}] = \frac{\beta}{2} \sum_{\mathbf{q}} \left(\frac{n_s e^2}{m} + \frac{q^2}{\mu_0} \right) (\mathbf{A}_{\mathbf{q}} \cdot \mathbf{A}_{-\mathbf{q}} - (\hat{\mathbf{q}} \cdot \mathbf{A}_{\mathbf{q}}) (\hat{\mathbf{q}} \cdot \mathbf{A}_{-\mathbf{q}})) \quad (10.95)$$

by defining $\mathbf{A}_{\mathbf{q}}^\perp \equiv \mathbf{A}_{\mathbf{q}} - \hat{\mathbf{q}}(\hat{\mathbf{q}} \cdot \mathbf{A}_{\mathbf{q}})$ and using the Coulomb Gauge to remove the $\mathbf{A}_{\mathbf{q}}^\parallel \equiv \hat{\mathbf{q}}(\hat{\mathbf{q}} \cdot \mathbf{A}_{\mathbf{q}})$ part, we can write

$$S[\mathbf{A}] = \frac{\beta}{2} \sum_{\mathbf{q}} \mathbf{A}_{\mathbf{q}}^\perp \left(\frac{n_s e^2}{m} + \frac{q^2}{\mu_0} \right) \mathbf{A}_{-\mathbf{q}}^\perp \quad (10.96)$$

see wee that the electromagnetic fields have gained mass, because the propagator is of the form $\mathcal{D}^{-1} = m^2 + |\mathbf{k}|^2$. This implies that there is attenuation of the electromagnetic field inside of superconductors, which is what leads to the Meissner Effect.

At the saddle point $\frac{\delta S}{\delta \mathbf{A}_{-\mathbf{q}}^\perp} = 0$ we have

$$\frac{\mu_0 n_s e^2}{m} \mathbf{A}_{\mathbf{q}}^\perp = -q^2 \mathbf{A}_{\mathbf{q}}^\perp \rightsquigarrow \nabla^2 \mathbf{B} = \lambda_L^{-2} \mathbf{B} \quad (10.97)$$

where $\lambda_L \equiv \sqrt{\frac{m}{\mu_0 n_s e^2}}$. This is the *First London Equation*, which we derived in CMP2 using Landau-Ginzburg Theory.

10.3.12 Are We Really Describing a Superconductor?

Our Free energy is of the form

$$F = v_F(\partial_\tau \theta - ie\phi)^2 + \frac{n_s}{2m}(\nabla\theta - e\mathbf{A})^2 + \frac{1}{\mu_0}(\nabla \times \mathbf{A})^2 \quad (10.98)$$

$\nabla_{\mathbf{A}}(\nabla \times \mathbf{A}) = 0$ because
 $(\nabla \times \mathbf{A})_x$ only depends on
 A_y and A_z .

From which we can calculate the current

$$\mathbf{j} = \nabla_{\mathbf{A}} F = \frac{n_s}{m}(\nabla\theta - e\mathbf{A}) \quad (10.99)$$

The derivative gives us

$$\frac{\partial \mathbf{j}}{\partial t} = \frac{en_s}{m} \mathbf{E} \quad (10.100)$$

telling us that for a constant electric field the current rises linearly as a function of time which is only possible if there is no resistance.

10.4 BOLTZMANN EQUATION

The Boltzmann Equation is an equation used in *kinetic theory*, which is the study of non-equilibrium phenomena in dilute gasses. In these dilute gasses particles follow Hamiltonian dynamics but suffer occasional scattering from other particles or impurities.

Jens uses $d\tau$, but we already
 integrate over something
 called τ in other sections,
 and I don't like using the
 same letters for different
 things.

The central object in kinetic theory is the phase-space distribution function $f(\mathbf{r}, \mathbf{p}, t)$. The quantity

$$f(\mathbf{r}, \mathbf{p}, t) \frac{d\mathbf{r}d\mathbf{p}}{(2\pi\hbar)^3} = f(\mathbf{r}, \mathbf{p}, t) d\mathbf{r}d\boldsymbol{\pi} \quad (10.101)$$

describes the mean number of particles in a phase-space volume of $d\mathbf{r}d\boldsymbol{\pi}$ at time t . We have introduced $d\boldsymbol{\pi}$ to avoid the factor of $(2\pi\hbar)^3$. In equilibrium the function $f(\mathbf{r}, \mathbf{p}, t)$ can be the Maxwell-Boltzmann distribution, the Fermi-Dirac distribution or the Bose-Einstein distribution or even something more exotic, depending on what the system is that you are describing.

The *Boltzmann Equation* is the continuity equation in phase space preserving the number of particles:

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} \cdot (\dot{\mathbf{r}}f) + \nabla_{\mathbf{k}} \cdot (\dot{\mathbf{k}}f) = 0 \quad (10.102)$$

this is the collisionless Boltzmann Equation, to include collisions we introduce the so-called *collision integral* on the right-hand side

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} \cdot (\dot{\mathbf{r}}f) + \nabla_{\mathbf{k}} \cdot (\dot{\mathbf{k}}f) = \left(\frac{\partial f}{\partial t} \right)_{\text{coll.}} \quad (10.103)$$

The collision integral is not a derivative, this is just notation; sometimes it is also written as \mathcal{I} . We should think of this as Hamiltonian evolution (lhs) with occasional collisions (rhs) that abruptly push you from one part of phase-space to another.

During the Hamiltonian evolution we have that $\nabla_{\mathbf{r}} \cdot \dot{\mathbf{r}} + \nabla_{\mathbf{k}} \cdot \dot{\mathbf{k}} = 0$ so

BOLTZMANN EQUATION

$$\frac{\partial f}{\partial t} + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}} f + \dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}} f = \left(\frac{\partial f}{\partial t} \right)_{\text{coll.}} \quad (10.104)$$

We can think of the Boltzmann Equation as the full-derivative of the distribution function f , which means that if we change variables we just need to adjust the names of the variables in the Boltzmann Equation, for instance if we want to change between canonical and kinetic momenta.

10.4.1 Collisions: Classical

Let us consider a classical scattering process in which we assume $d\mathbf{r}$ is fixed. The probability of scattering during the time interval dt is $d\pi w(\mathbf{p}', \mathbf{p}) dt$, so the rate of change in f due to collisions is

$$\left(\frac{\partial f}{\partial t} \right)_{\text{coll.}} = - \int d\pi [w(\mathbf{p}', \mathbf{p}) f(\mathbf{p}) - w(\mathbf{p}, \mathbf{p}') f(\mathbf{p}')] \quad (10.105)$$

The 2 particle version thereof would be

$$\begin{aligned} \left(\frac{\partial f_1}{\partial t} \right)_{\text{coll.}} = & - \int d\pi_2 d\pi'_1 d\pi'_2 [w(\mathbf{p}'_1, \mathbf{p}'_2; \mathbf{p}_1, \mathbf{p}_2) f(\mathbf{p}_1) f(\mathbf{p}_2) \\ & - w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) f(\mathbf{p}'_1) f(\mathbf{p}'_2)] \end{aligned} \quad (10.106)$$

where we've assumed that the two-particle distribution function splits into the product of two one-particle distribution functions, which is valid in the low-density limit ($f(\mathbf{p}_1; \mathbf{p}_2) = f(\mathbf{p}_1) f(\mathbf{p}_2)$).

For time-reversal symmetric systems we have that

$$w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) = w(-\mathbf{p}'_1, -\mathbf{p}'_2; -\mathbf{p}_1, -\mathbf{p}_2) \quad (10.107)$$

and for inversion-symmetric systems we have

$$w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) = w(-\mathbf{p}_1, -\mathbf{p}_2; -\mathbf{p}'_1, -\mathbf{p}'_2) \quad (10.108)$$

if we have a system that is both time-reversal and inversion symmetric then it must hold that

$$w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) = w(\mathbf{p}'_1, \mathbf{p}'_2; \mathbf{p}_1, \mathbf{p}_2) \quad (10.109)$$

which is referred to as the *Principle of Detailed Balance*. In such systems the collision integral simplifies to

$$\left(\frac{\partial f_1}{\partial t} \right)_{\text{coll.}} = \int d\pi_2 d\pi'_1 d\pi'_2 w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) [f'_1 f'_2 - f_1 f_2] \quad (10.110)$$

where I have omitted the \mathbf{p} s.

10.4.2 Boltzmann H-Theorem

Let us introduce

$$\rho_\phi(\mathbf{r}, t) = \int d\pi \phi(\mathbf{p}, f) f(\mathbf{r}, \mathbf{p}, t) \quad (10.111)$$

for some function(al) ϕ . Its rate of change due to 2-body collisions is

$$\left(\frac{\partial \rho_\phi}{\partial t}\right)_{\text{coll.}} = \int d\pi_1 \pi_2 \pi'_1 \pi'_2 \bar{\phi} w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) [f'_1 f'_2 - f_1 f_2] \quad (10.112)$$

where we have defined $\bar{\phi} \equiv \phi + f \frac{\partial \phi}{\partial f}$. Manipulating the w term and using the principle of detailed balance we obtain

$$\left(\frac{\partial \rho_\phi}{\partial t}\right)_{\text{coll.}} = -\frac{1}{4} \int d\pi_1 \pi_2 \pi'_1 \pi'_2 w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) (\bar{\phi}_1 + \bar{\phi}_2 - \bar{\phi}'_1 - \bar{\phi}'_2) (f_1 f_2 - f'_1 f'_2) \quad (10.113)$$

suppose $\phi = \ln f$

$$\left(\frac{\partial}{\partial t} \int d\pi_1 f_1 \ln f_1\right)_{\text{coll.}} = -\frac{1}{4} \int d\pi_1 \pi_2 \pi'_1 \pi'_2 w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) (f_1 f_2 - f'_1 f'_2) \ln \left(\frac{f_1 f_2}{f'_1 f'_2}\right) \quad (10.114)$$

the right-hand side is always negative, which means that $H = \int d\pi_1 f_1 \ln f_1$ (enthalpy) always decreases due to collisions. Relating enthalpy to entropy tells us that entropy always increases due to collisions.

10.4.3 Conservation Laws

If ϕ is chosen to be independent of f then the density

$$\rho_\phi(\mathbf{r}, t) = \int d\pi \phi(\mathbf{p}) f(\mathbf{r}, \mathbf{p}, t) \quad (10.115)$$

is unchanged by collisions if ϕ is a collision invariant (because then $\phi_1 + \phi_2 = \phi'_1 + \phi'_2$). We have a number of invariants: energy, E , momentum, $\pm p$, and numbers. When we minimise the enthalpy we require $f_1 f_2 = f'_1 f'_2$, so $\ln f_1 + \ln f_2 = \ln f'_1 + \ln f'_2$. So $\ln f$ is also a collision invariant, which implies we can express it as a linear combination of the aforementioned collision invariants:

$$\ln f = A + \mathbf{B} \cdot \mathbf{p} + CE \quad (10.116)$$

suppose $A = \beta\mu$, $\mathbf{B} = \beta\mathbf{u}$ (drift velocity) and $C = -\beta$, we get

$$f = \exp[-\beta(E - \mathbf{p} \cdot \mathbf{u} - \mu)] \quad (10.117)$$

10.4.4 Collisions: Quantum

In the quantum mechanical Boltzmann Equation we need to account for Pauli's Exclusion principle (for fermions). For bosons we acquire a similar factor which accounts for Bose-statistics which makes it seem like bosons "prefer" to be in same states. Thus we need to modify the $f_1 f_2 - f'_1 f'_2$ term in the Boltzmann Equation:

$$f_1 f_2 - f'_1 f'_2 \rightsquigarrow f_1 f_2 (1 + \zeta f'_1) (1 + \zeta f'_2) - f'_1 f'_2 (1 + \zeta f_1) (1 + \zeta f_2) \quad (10.118)$$

where $\zeta = 1$ for bosons and $\zeta = -1$ for fermions.

In equilibrium this vanishes, which gives us

$$\ln\left(\frac{f_1}{1+\zeta f_1}\right) + \ln\left(\frac{f_2}{1+\zeta f_2}\right) = \ln\left(\frac{f'_1}{1+\zeta f'_1}\right) + \ln\left(\frac{f'_2}{1+\zeta f'_2}\right) \quad (10.119)$$

So if $\ln\left(\frac{f}{1+\zeta f}\right) = x = \beta(E - \mathbf{p} \cdot \mathbf{u} - \mu)$ we get

$$\frac{f}{1+\zeta f} = e^x \rightsquigarrow f = \frac{1}{e^{-x} - \zeta} \quad (10.120)$$

giving us both the Bose-Einstein and the Fermi-Dirac distribution functions.

10.4.5 Linearisation

Consider the collision integral from Equation 10.110:

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll.}} = \int d\pi_2 d\pi'_1 d\pi'_2 w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) [f'_1 f'_2 - f_1 f_2] \quad (10.121)$$

Now let us assume that we are close to equilibrium, such that we can write $f = f^0 + f^0 \psi$ for $\psi \ll 1$, where f^0 is the equilibrium distribution function. We have that $f_1^0 f_2^0 = f_1'^0 f_2'^0$ (the collision integral is zero in equilibrium), and means that up to first order in ψ we can write

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll.}} = \int d\pi_2 d\pi'_1 d\pi'_2 w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) f_1^0 [\psi'_1 + \psi'_2 - \psi_1 - \psi_2] \quad (10.122)$$

Additionally splitting the left-hand side up and using that f_0 is not a function of time we get that

$$\left(\frac{\partial \psi_1}{\partial t}\right)_{\text{coll.}} = \int d\pi_2 d\pi'_1 d\pi'_2 w(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}'_1, \mathbf{p}'_2) [\psi'_1 + \psi'_2 - \psi_1 - \psi_2] \quad (10.123)$$

So we have linearised the collision integral. Because the Boltzmann Equation is now a linear integro-differential equation, there is a lot of theory that we learned about in Quantum Mechanics that can be used. For instance there exists a basis of functions that span the general solution to the integro-differential equation. However, keep in mind that this only holds for small values of ψ .

10.4.6 Relaxation-Time Approximation

In the relaxation time approximation we assume that the collision integral is given by

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll.}} = -\frac{f - f_0}{\tau} \quad (10.124)$$

*Jens Paaske's note here is
"One must be careful not to
break conservation laws!"*

where τ is the relaxation time of the system. If you start a system at some $f \neq f_0$ then you can expect to wait $\sim \tau$ before the system returns to equilibrium.

Consider now the Boltzmann Equation where we ignore the $\left(\frac{\partial f}{\partial t}\right)$ and $\nabla_{\mathbf{k}}$ terms and use the relaxation-time approximation:

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f = -\frac{f - f_0}{\tau} \quad (10.125)$$

assume that f^0 only depends on $\beta(E - \mu)$, which means that the left-hand side becomes

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f \approx \mathbf{v} \cdot \nabla_{\mathbf{r}} f^0 = -\frac{\partial f^0}{\partial E} \mathbf{v} \cdot (E \beta \nabla T - \mu \beta \nabla T + \nabla \mu) \quad (10.126)$$

the last two terms can be written in terms of the enthalpy per particle, h , reducing it to

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f \approx f^0 (\mathbf{v} \cdot \nabla T) k_B \beta^2 (E - h) \quad (10.127)$$

whereas the collision integral becomes

$$-\frac{f - f^0}{\tau} = -\frac{f^0 \psi}{\tau} \quad (10.128)$$

giving us

$$\psi = -k_B \beta^2 \tau (E - h) \mathbf{v} \cdot \nabla T \quad (10.129)$$

using which we can calculate the heat current due to the gradient in the temperature

$$\mathbf{j}_e = \int \pi \mathbf{v} (\mathbf{v} \cdot \nabla T) k_B \beta^2 \tau (h - E) f^0 E = -\kappa \nabla T \quad (10.130)$$

where we can evaluate the thermal conductivity

$$\kappa = \int d\pi \frac{v^2}{3} E \tau (E - h) f^0 \beta = \frac{5 n k_B^2 T}{2m} \quad (10.131)$$

(the final equality holds if we use the Maxwell Distribution for f^0)

10.5 NON-EQUILIBRIUM KELDYSH FIELD THEORY

10.5.1 Closed Time Contour

Despite the fact that non-equilibrium Keldysh Field Theory considers real-time values, it is general enough to encompass descriptions of non-zero temperature systems.

We consider a system whose state at $t = -\infty$ is *known* and is given by the density matrix $\hat{\rho}(-\infty)$. We assume that the Hamiltonian is time-dependent and that $\hat{H}(-\infty)$ describes a system of non-interacting particles. At finite times the interactions are turned on adiabatically, such that they have reached their actual value before we perform our “measurement”. The density matrix evolves according to the Von Neumann Equation

$$\partial_t \hat{\rho}(t) = -i[\hat{H}(t), \hat{\rho}(t)] \quad (10.132)$$

because we know the state of system at $t = -\infty$ and we can calculate the time-evolution operator

$$\hat{U}_{t,t'} = \mathbb{T} \exp \left(-i \int_{t'}^t dt'' H(t'') \right) \quad (10.133)$$

we can solve the Von Neumann Equation by simply evolving the state

$$\hat{\rho}(t) = \hat{U}_{t,-\infty} \hat{\rho}(-\infty) \hat{U}_{t,-\infty}^\dagger \quad (10.134)$$

We are often interested in expectation values of an operator, \hat{O} :

$$\langle O \rangle(t) \equiv \frac{\text{tr}[\hat{O}\hat{\rho}(t)]}{\text{tr}[\hat{\rho}(t)]} = \frac{\text{tr}[\hat{U}_{-\infty,t}\hat{O}\hat{U}_{t,-\infty}\hat{\rho}(-\infty)]}{\text{tr}[\hat{\rho}(t)]} \quad (10.135)$$

where we have used the cyclicity of the trace and that $\hat{U}_{t,t'}^\dagger = \hat{U}_{t',t}$. This is also the starting point of equilibrium field theory, however, there one assumes that $\hat{U}_{\infty,\infty}|0\rangle = e^{iL}|0\rangle$, which is what leads to the Gell-Mann-Low Formula see [coleman].

For non-equilibrium field theory, however, we proceed by inserting $\mathbb{1} = \hat{U}_{t,+\infty}\hat{U}_{+\infty,t}$ just to the left of \hat{O} such that we obtain

$$\langle O \rangle(t) = \frac{\text{tr}[\hat{U}_{-\infty,+\infty}\hat{U}_{+\infty,t}\hat{O}\hat{U}_{t,-\infty}\hat{\rho}(-\infty)]}{\text{tr}[\hat{\rho}(t)]} \quad (10.136)$$

This describes the evolution along the *closed time contour*, C , which we refer to as the *Keldysh Contour*. Notice how we could have inserted the $\mathbb{1}$ just to the *right* of \hat{O} , which would have put the observable on the backward branch of the contour, see Figure 10.5.

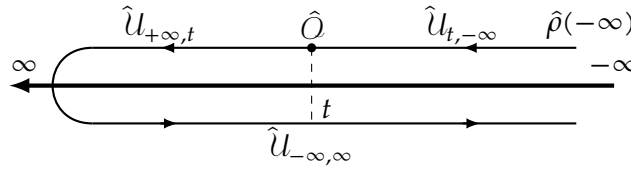


Figure 10.4: The Keldysh Contour, C [ak]

In Keldysh Field Theory it is customary to define the partition function as

$$Z = \langle \mathbb{1} \rangle = \frac{\text{tr}[\hat{U}_C\hat{\rho}(-\infty)]}{\text{tr}[\hat{\rho}(t)]} = 1 \quad (10.137)$$

where we have defined $\hat{U}_C = \hat{U}_{-\infty,+\infty}\hat{U}_{+\infty,\infty}$ which is unity if the Hamiltonian is the same on the forward and backward portion of the contour. Thus we will often work with $Z = 1$, which means that we cannot derive physical quantities directly from it, as one would in statistical physics; despite this we will gain a rich understand of non-equilibrium field theory.

Let us define $\hat{H}_V^\pm(t) = \hat{H}(t) \pm \hat{O}V(t)$ where the plus (minus) means that we use this expression for \hat{H}_V on the forward (backward) branch. Since we have constructed a Hamiltonian that is not the same on the two branches our $Z[V]$ becomes a non-trivial functional of V :

$$Z[V] = \frac{\text{tr}[\hat{U}_C[V]\hat{\rho}(-\infty)]}{\text{tr}[\hat{\rho}(t)]} \quad (10.138)$$

this is the *generating functional*. From here we can obtain

$$\langle \hat{O} \rangle(t) = \left. \frac{i}{2} \frac{\delta Z[V]}{\delta V(t)} \right|_{V=0} \quad (10.139)$$

which is reminiscent of the corresponding expression in equilibrium many-body physics.

10.5.2 Bosons

Let us begin with a single bosonic Harmonic Oscillator

$$\hat{H} = \omega_0 \hat{b}^\dagger \hat{b} \quad (10.140)$$

and let us assume that the initial density matrix was

$$\hat{\rho}_{-\infty} = \exp(-\beta(\hat{H} - \mu\hat{N})) = \exp(-\beta(\omega_0 - \mu)\hat{N}) \quad (10.141)$$

Our Hamiltonian is the same on both parts of the contour, so $Z = 1$:

$$Z = \frac{\text{tr}[\hat{\mathcal{U}}_C \hat{\rho}_{-\infty}]}{\text{tr}[\hat{\rho}_{-\infty}]} \quad (10.142)$$

let us calculate the denominator

$$\begin{aligned} \text{tr}[\hat{\rho}_{-\infty}] &= \sum_{n=0}^{\infty} e^{-\beta(\omega_0 - \mu)n} = [1 - \rho(\omega_0)]^{-1}, \\ \hat{\rho}(\omega_0) &= \exp[-\beta(\omega_0 - \mu)] \end{aligned} \quad (10.143)$$

let us discretise the Keldysh Contour that we can calculate the numerator by hand:

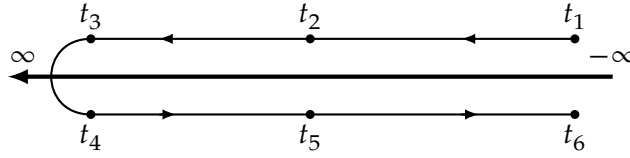


Figure 10.5: A Discretised Keldysh Contour, $C[\mathbf{ak}]$

Define $\Delta t = t_{j+1} - t_j$, and now expand the numerator

$$\begin{aligned} \text{tr}[\hat{\mathcal{U}}_C \hat{\rho}_{-\infty}] &\sim \langle \phi_6 | \hat{\mathcal{U}}_{-\Delta t} | \phi_5 \rangle \langle \phi_5 | \hat{\mathcal{U}}_{-\Delta t} | \phi_4 \rangle \langle \phi_4 | \mathbb{1} | \phi_3 \rangle \times \\ &\quad \langle \phi_3 | \hat{\mathcal{U}}_{\Delta t} | \phi_2 \rangle \langle \phi_2 | \hat{\mathcal{U}}_{\Delta t} | \phi_1 \rangle \langle \phi_1 | \hat{\rho}_{-\infty} | \phi_6 \rangle \end{aligned} \quad (10.144)$$

We are considering coherent states, so as long as the Hamiltonian is normal ordered we can use

$$\langle \phi_j | \hat{\mathcal{U}}_{\pm \Delta t} | \phi_{j-1} \rangle \approx e^{\bar{\phi}_j \phi_{j-1}} e^{\mp i H[\bar{\phi}_j, \phi_{j-1}] \Delta} \quad (10.145)$$

where we've assumed Δt is sufficiently small that we can use

$$\begin{aligned} \langle \phi_i | \exp(a\hat{H}) | \phi_j \rangle &\approx \langle \phi_i | 1 + H[\bar{\phi}_i, \phi_j] | \phi_j \rangle \\ &\approx \langle \phi_i | \exp(a\hat{H}[\bar{\phi}_i, \phi_j]) | \phi_j \rangle \end{aligned} \quad (10.146)$$

The part containing $\hat{\rho}_{-\infty}$:

$$\langle \phi_1 | \hat{\rho}_{-\infty} | \phi_6 \rangle = e^{\bar{\phi}_1 \phi_6 \rho(\omega_0)} \quad (10.147)$$

Resulting in

$$Z = [1 - \rho(\omega_0)] \int \prod_{j=1}^{2N} d[\bar{\phi}_j, \phi_j] \exp(i\bar{\phi} \cdot \underline{G}^{-1} \cdot \phi) \quad (10.148)$$

where in our case $N = 3$ and

$$i\mathcal{G}^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & \rho(\omega_0) \\ h_- & -1 & 0 & 0 & 0 & 0 \\ 0 & h_- & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & h_+ & -1 & 0 \\ 0 & 0 & 0 & 0 & h_+ & -1 \end{pmatrix} \quad (10.149)$$

and we have defined $h_{\pm} \equiv 1 \pm i\omega_0\Delta t$. We can evaluate the Gaussian integral giving us

$$\frac{1}{\det(-i\mathcal{G}^{-1})} = \frac{1}{1 - \rho(\omega_0)(h_-h_+)^{N-1}} \xrightarrow{N \rightarrow \infty} \frac{1}{1 - \rho(\omega_0)} \quad (10.150)$$

so we see that $Z = 1$ once we take the $N \rightarrow \infty$ limit. Let us write things more explicitly

$$Z = \int \mathcal{D}[\bar{\phi}(t), \phi(t)] \exp(iS[\bar{\phi}, \phi]) \quad (10.151)$$

where $\mathcal{D}[\bar{\phi}(t), \phi(t)] = \frac{1}{\text{tr} \bar{\rho}_{-\infty}} \prod_{j=1}^{2N} d[\bar{\phi}_j, \phi_j]$ and the action is given by

$$S[\bar{\phi}, \phi] = \sum_{j=2}^{2N} \delta t_j \left[i\bar{\phi}_j \left(\frac{\phi_j - \phi_{j-1}}{\delta t_j} \right) - \omega_0 \bar{\phi}_j \phi_{j-1} \right] + i\bar{\phi}_1[\phi_1 - \rho(\omega_0)\phi_{2N}] + \quad (10.152)$$

$$i\bar{\phi}_{N+1}[\phi_{N+1} - \phi_N]$$

where we've defined $\delta t_j = \pm(t_j - t_{j-1})$ (the \pm depends on which branch of the contour we are on). Notice the similarity between the first term and Equation 10.19, however, the remaining terms were not seen in our previous formalism. Ignoring the latter terms we can write the first term in the action as

$$S_G[\bar{\phi}, \phi] = \oint_C dt \bar{\phi}(t) G^{-1} \phi(t), \quad G^{-1} = i\partial_t - \omega_0 \quad (10.153)$$

We now split $\phi(t)$ up into its forward-branch part (ϕ^+) and its backward-branch part (ϕ^-), which means we can write

$$S_G[\bar{\phi}, \phi] = \int_{-\infty}^{\infty} dt \begin{pmatrix} \bar{\phi}^+(t) & \bar{\phi}^-(t) \end{pmatrix} \tau_3 (i\partial_t - \omega_0) \begin{pmatrix} \phi^+(t) \\ \phi^-(t) \end{pmatrix} \quad (10.154)$$

as it stands ϕ^{\pm} are completely independent, but in reality they are linked by the latter terms in Equation 10.153.

Green's Functions

Let us now calculate Green's functions, which we can do by inverting Equation 10.149:

$$i\mathcal{G} = \frac{1}{g} \begin{pmatrix} 1 & h_- h_+^2 \rho & h_+^2 \rho & h_+^2 \rho & h_+ \rho & \rho \\ h_- & 1 & h_- h_+^2 \rho & h_- h_+^2 \rho & h_- h_+ \rho & h_- \rho \\ h_-^2 & h_- & 1 & h_-^2 h_+^2 \rho & h_-^2 h_+ \rho & h_-^2 \rho \\ h_-^2 & h_- & 1 & 1 & h_-^2 h_+ \rho & h_-^2 \rho \\ h_-^2 h_+ & h_- h_+ & h_+ & h_+ & 1 & h_-^2 h_+ \rho \\ h_-^2 h_+^2 & h_- h_+^2 & h_+^2 & h_+^2 & h_+ & 1 \end{pmatrix} \quad (10.155)$$

where $g \equiv \det[-i\underline{G}^{-1}]$ and $\rho \equiv \rho(\omega_0)$. These can be split up into the lesser than, greater than, time-ordered and anti time-ordered Green's functions, that evaluate to

$$\begin{aligned}\langle \phi^+(t) \bar{\phi}^-(t') \rangle &= iG_{t,t'}^< = n_B(\omega_0) \exp(-i\omega_0(t-t')) \\ \langle \phi^-(t) \bar{\phi}^+(t') \rangle &= iG_{t,t'}^> = (1 + n_B(\omega_0)) \exp(-i\omega_0(t-t')) \\ \langle \phi^+(t) \bar{\phi}^+(t') \rangle &= iG_{t,t'}^{\mathbb{T}} = i\theta(t-t')G_{t,t'}^> + i\theta(t'-t)G_{t,t'}^< \\ \langle \phi^-(t) \bar{\phi}^-(t') \rangle &= iG_{t,t'}^{\bar{\mathbb{T}}} = i\theta(t-t')G_{t,t'}^< + i\theta(t'-t)G_{t,t'}^>\end{aligned}\quad (10.156)$$

Remember we are still looking at the single bosonic level in “thermal equilibrium” (in the sense that its occupations are given by the Bose-Einstein distribution)

Clearly these four functions are interrelated, which can also be seen in the following

$$G_{t,t'}^{\mathbb{T}} + G_{t,t'}^{\bar{\mathbb{T}}} - G_{t,t'}^> - G_{t,t'}^< = -i\delta_{t,t'} \quad (10.157)$$

hence there are only three independent Green's functions. This is seen by changing basis; we perform a *Keldysh Rotation*, which is a unitary transformation:

$$\begin{pmatrix} \phi^{\text{cl}} \\ \phi^{\text{q}} \end{pmatrix} = \frac{1}{\sqrt{2}}(\tau_1 + \tau_3) \begin{pmatrix} \phi^+ \\ \phi^- \end{pmatrix} \quad (10.158)$$

The transformed Green's matrix is

$$i\underline{G}_{t,t'} = \begin{pmatrix} iG_{t,t'}^{\text{K}} & iG_{t,t'}^{\text{R}} \\ iG_{t,t'}^{\text{A}} & 0 \end{pmatrix} \quad (10.159)$$

where at $t \neq t'$ it holds that

$$G^{\text{K}} = G^< + G^>, \quad G^{\text{R}} = \theta(t-t')(G^> - G^<), \quad G^{\text{A}} = \theta(t'-t)(G^< - G^>) \quad (10.160)$$

Fourier transforming

$$\begin{aligned}G^{\text{K}}(\omega) &= -2\pi i[2n_B(\omega) + 1]\delta(\omega - \omega_0), & G^{\text{R}}(\omega) &= \frac{1}{\omega - \omega_0 + i0^+}, \\ G^{\text{A}}(\omega) &= \frac{1}{\omega - \omega_0 - i0^+}\end{aligned} \quad (10.161)$$

So in equilibrium we get

$$G^{\text{K}}(\epsilon) = \coth\left(\frac{\epsilon - \mu}{2T}\right)[G^{\text{R}}(\epsilon) - G^{\text{A}}(\epsilon)] \quad (10.162)$$

This result also holds for other systems in *thermal equilibrium*.

It is important here to recognise the retarded and advanced Green's functions that we've known about since CMT1 (although we didn't call them that in the course), and to see that this new Green's function, the Keldysh Green's function, carries information about occupations. A common Ansatz for G^{K} is

$$G^{\text{K}} = G^{\text{R}} \circ F - F \circ G^{\text{A}}, \quad (F \circ G)_{t,t'} \equiv \int_{-\infty}^{\infty} dt'' F(t, t'')G(t'', t') \quad (10.163)$$

where F is the distribution function (equilibrium or non-equilibrium) – in the equilibrium case we had $F(\omega) = \coth\left(\frac{\omega - \mu}{2T}\right)$.

10.5.3 Keldysh Action

The inverse of the matrix in Equation 10.159 has a zero $(0, 0)$ -component, our notation in the course was

$$G^{-1} = \begin{pmatrix} 0 & G_A^{-1} \\ G_R^{-1} & G_K^{-1} \end{pmatrix} \quad (10.164)$$

so the action is

$$S[\phi^{\text{cl}}, \phi^{\text{q}}] = \iint_{-\infty}^{\infty} dt dt' \begin{pmatrix} \bar{\phi}^{\text{cl}} & \bar{\phi}^{\text{q}} \end{pmatrix}_t \begin{pmatrix} 0 & G_A^{-1} \\ G_R^{-1} & G_K^{-1} \end{pmatrix}_{tt'} \begin{pmatrix} \phi^{\text{cl}} \\ \phi^{\text{q}} \end{pmatrix}_{t'} \quad (10.165)$$

we see that if $\phi^{\text{q}} = 0$ then $S = 0$, which is equivalent to saying that if H is the same on both branches of the contour $Z = 1$.

We have that

$$[G_{R(A)}^{-1}]_{t,t'} = \begin{cases} \delta(t-t')(i\vec{\partial}_{t'} - \omega_0 \pm i0^+) & \text{to the right} \\ \delta(t-t')(-i\vec{\partial}_t - \omega_0 \pm i0^+) & \text{to the left} \end{cases} \quad (10.166)$$

And we have that

$$G_K^{-1} = G_R^{-1} \circ F - F \circ G_A^{-1} \quad (10.167)$$

which means that in equilibrium we can write

$$G^{-1} = \begin{pmatrix} 0 & \omega - \omega_0 - i0^+ \\ \omega - \omega + i0^+ & 2i0^+ \coth\left(\frac{\omega - \mu}{2T}\right) \end{pmatrix} \quad (10.168)$$

10.5.4 Feynman Diagrams

The Keldysh, retarded and advanced Green's functions can be represented in terms of Feynman Diagrams, in this course we used

$$G^K = \longrightarrow$$

$$G^R = \longrightarrow \dots$$

$$G^A = \dots \longrightarrow \quad (10.169)$$

10.5.5 From Keldysh to Langevin

Let us begin with the action for the single boson

$$S[\bar{\phi}, \phi] = \oint_C dt \bar{\phi}(t) G^{-1} \phi(t), \quad G^{-1} = i\partial_t - \omega_0 \quad (10.170)$$

and let us rewrite the complex fields in terms of their real and imaginary parts

$$\phi(t) = \frac{1}{\sqrt{2}\omega_0} (p(t) - i\omega_0 x(t)) \quad (10.171)$$

for $x(t), p(t) \in \mathbb{R}$. Putting this in for ϕ (and the conjugate for $\bar{\phi}$) and multiplying out we get

$$S[x, p] = \oint_C dt \left[p\dot{x} - \frac{1}{2}p^2 - \frac{1}{2}\omega_0^2 x^2 \right] + \text{const.} \quad (10.172)$$

where const. is the integral of a total derivative and hence doesn't affect the equations of motion. Our partition function becomes

$$Z = \int \mathcal{D}[x, p] \exp \left(i \oint_C (p\dot{x} - H) \right) \quad (10.173)$$

notice how it is the Lagrangian we are integrating over – the Legendre transform of the Hamiltonian. Hence, generally we could write

$$Z = \int \mathcal{D}[x, p] \exp \left(i \oint_C (p\dot{x} - \frac{1}{2}p^2 - V(x)) \right) \quad (10.174)$$

or, if we integrate out the momentum

$$S[x] = \oint_C dt \left(\frac{1}{2}\dot{x}^2 - V(x) \right) \quad (10.175)$$

However, let us now split the contour into the forward and backwards branch and change variables to the classical and quantum fields:

$$S[x^{\text{cl}}, x^{\text{q}}] = \int_{-\infty}^{\infty} dt \left(-2x^{\text{q}}\ddot{x}^{\text{cl}} - V(x^{\text{cl}} + x^{\text{q}}) + V(x^{\text{cl}} - x^{\text{q}}) \right) \quad (10.176)$$

Suppose now we can treat x^{q} as small variables, such that we only need to expand the potential up to linear order in it:

$$S[x^{\text{cl}}, x^{\text{q}}] = \int_{-\infty}^{\infty} dt \, 2x^{\text{q}} (\ddot{x}^{\text{cl}} + V'(x^{\text{cl}})) \quad (10.177)$$

it turns out that $\int \mathcal{D}[x] e^{i \int dt \, xf} = \delta(f)$ which means integrating out x^{q} gives us Newton's Second Law:

$$\ddot{x}^{\text{cl}} = -V'(x^{\text{cl}}) \quad (10.178)$$

However, if x^{q} isn't a small parametre our formalism can account for the quantum fluctuations.

Coupling the Harmonic Oscillator to a Bath

Let us now include an interaction between the single boson level with a bath

$$\begin{aligned} S_{\text{HO}}[x] &= \int_{-\infty}^{\infty} dt \left(-2x^{\text{q}}\ddot{x}^{\text{cl}} - V(x^{\text{cl}} + x^{\text{q}}) + V(x^{\text{cl}} - x^{\text{q}}) \right) \\ S_{\text{bath}}[\phi] &= \frac{1}{2} \sum_s \int_{-\infty}^{\infty} dt \, \bar{\phi}_s^T D_s^{-1} \phi_s \\ S_{\text{int}}[x, \phi] &= \sum_s g_s \int_{-\infty}^{\infty} dt \, \bar{x}^T \tau_1 \phi_s \end{aligned} \quad (10.179)$$

Let us integrate out the bath to obtain the dissipative action

$$S_{\text{diss}} = \frac{1}{2} \iint_{-\infty}^{\infty} dt dt' \tilde{x}^T(t) \mathbb{D}^{-1}(t-t') x(t') \quad (10.180)$$

where

$$\mathbb{D}^{-1}(t-t') = -\tau_1 \left[\sum_s g_s^2 D_s(t-t') \right] \tau_1 \quad (10.181)$$

\mathbb{D} has the same “causality structure” as we would expect from a Green’s function, that is to say that it makes sense to assign a retarded, advanced and Keldysh component to it and that they all behave as they should. The retards and advanced components are

$$\mathbb{D}_{\text{R/A}}^{-1} = \frac{1}{2} \sum_s \frac{g_s^2}{(\epsilon \pm i0)^2 - \omega_s^2} = \int_0^{\infty} \frac{d\omega}{2\pi} \frac{\omega J(\omega)}{\omega^2 - (\epsilon \pm i0)^2} \quad (10.182)$$

where we’ve defined $J(\omega) \equiv \pi \sum_s \frac{g_s^2}{\omega_s} \delta(\omega - \omega_s)$. If we assume that $J(\omega) = 8\gamma\omega$ where γ is constant at small frequencies we find that

$$\begin{aligned} \mathbb{D}_{\text{R/A}}^{-1}(\epsilon) &= 4\gamma \int_0^{\infty} \frac{d\omega}{2\pi} \frac{\omega^2}{\omega^2 - (\epsilon \pm i0)^2} \\ &= \text{const.} \pm 2i\gamma\epsilon \end{aligned} \quad (10.183)$$

which means that the Keldysh component is

$$\mathbb{D}_{\text{K}}^{-1} = 4i\gamma\epsilon \coth\left(\frac{\epsilon}{2T}\right) \quad (10.184)$$

Thus we have a non-zero imaginary part now, which means there are non-zero lifetimes too. In t -space we get

$$\mathbb{D}_{\text{K}}^{-1}(t-t') = 4i\gamma \left[(2T + C)\delta(t-t') - \frac{\pi T^2}{\sinh^2(\pi T(t-t'))} \right] \quad (10.185)$$

and $\mathbb{D}_{\text{R/A}}^{-1} = \mp\gamma\delta(t-t')\partial_{t'}$. The resulting action is

$$\begin{aligned} S[x] &= \int_{-\infty}^{\infty} dt \left[-2x^{\text{q}}(\ddot{x}^{\text{cl}} + \gamma\dot{x}^{\text{cl}}) - V(x^{\text{cl}} + x^{\text{q}}) + V(x^{\text{cl}} - x^{\text{q}}) \right] \\ &+ 2i\gamma \int_{-\infty}^{\infty} dt \left[2T(x^{\text{q}}(t))^2 + \frac{\pi T^2}{2} \int_{-\infty}^{\infty} dt' \frac{[x^{\text{q}}(t) - x^{\text{q}}(t')]^2}{\sinh^2[\pi T(t-t')]} \right] \end{aligned} \quad (10.186)$$

Now it is time to reinstate \hbar , we do so by dividing the entire action by \hbar (so that $\exp(S/\hbar)$ has a unitless argument), but we also let $x^{\text{q}} \rightarrow \hbar x^{\text{q}}$ and $T \rightarrow T/\hbar$. Thus all linear terms in x^{q} aren’t affected, we expand $V(x^{\text{cl}} \pm x^{\text{q}})$ up to linear order in x^{q} and we need to treat the last term separately when we let $\hbar \rightarrow 0$. We get

$$\lim_{\hbar \rightarrow 0} \frac{\pi T^2}{2\hbar \sinh^2(\pi T(t-t')/\hbar)} = T\delta(t-t') \quad (10.187)$$

which implies that the last term falls out completely leaving only

$$S[x] = \int_{-\infty}^{\infty} dt \left[-2x^{\text{q}}[\ddot{x}^{\text{cl}} + \gamma\dot{x}^{\text{cl}} + V'(x^{\text{cl}})] + 4i\gamma T[x^{\text{q}}(t)]^2 \right] \quad (10.188)$$

we are almost ready to integrate x^q out again, but first we need to deal with the quadratic term. We do so by means of a Hubbard-Stratonovich transformation:

$$e^{-4\gamma T \int dt [x^q(t)]^2} = \int \mathcal{D}[\xi(t)] e^{-\int dt [\frac{1}{4\gamma T} \xi^2(t) - 2i\xi(t)x^q(t)]} \quad (10.189)$$

where the measure $\mathcal{D}[\xi(t)]$ is chosen such that $\int \mathcal{D}[\xi(t)] e^{-\int dt [\frac{1}{4\gamma T} \xi^2(t)]} = 1$. The expectation value of an observable $O[x^{\text{cl}}]$ can be written as

$$\langle O \rangle = \int \mathcal{D}[x^q, x^{\text{cl}}] O[x^{\text{cl}}] e^{iS[x]} = \int \mathcal{D}[\xi, x^{\text{cl}}] e^{-\int dt \frac{\xi^2}{4\gamma T}} O[x^{\text{cl}}] \delta(\ddot{x}^{\text{cl}} + \gamma \dot{x}^{\text{cl}} + V'(x^{\text{cl}}) - \xi) \quad (10.190)$$

hence the only functions (trajectories) x^{cl} that contribute to $\langle \hat{O} \rangle$ are the ones that satisfy the *Langevin Equation*

$$\ddot{x}^{\text{cl}} = -\gamma \dot{x}^{\text{cl}} - V'(x^{\text{cl}}) + \xi(t) \quad (10.191)$$

which is Newton's Second Law with friction and with a *random* driving force, for which

$$\langle \xi(t) \xi(t') \rangle = \int \mathcal{D}[\xi] \xi(t) \xi(t') e^{-\int dt \frac{\xi^2}{4\gamma T}} = 2\gamma T \delta(t - t') \quad (10.192)$$

random forces of this type are often referred to as *white noise forces*.

10.5.6 From Keldysh to the Boltzmann Equation

Let us now begin by extending the Keldysh formalism to multiple bosons, *i.e* now

$$H_0 = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \hat{b}_{\mathbf{k}}^\dagger \hat{b}_{\mathbf{k}} \quad (10.193)$$

where our initial density matrix is $\hat{\rho}_{-\infty} = \exp[-\beta(\hat{H}_0 - \mu \hat{N})]$ and $\hat{N} \equiv \sum_{\mathbf{k}} \hat{b}_{\mathbf{k}}^\dagger \hat{b}_{\mathbf{k}}$. Similarly to before we get

$$Z = \int \mathcal{D}[\phi^{\text{cl}}, \phi^q] e^{iS_0[\phi^{\text{cl}}, \phi^q]} = 1 \quad (10.194)$$

where

$$\mathcal{D}[\phi^{\text{cl}}, \phi^q] = \frac{1}{\text{tr}[\rho_{-\infty}]} \prod_{\mathbf{k}} \prod_{j=1}^N \frac{d\text{Re}\phi_j^{\text{cl}}(\mathbf{k}) d\text{Im}\phi_j^{\text{cl}}(\mathbf{k})}{\pi} \frac{d\text{Re}\phi_j^q(\mathbf{k}) d\text{Im}\phi_j^q(\mathbf{k})}{\pi} \quad (10.195)$$

And now in the action we sum over momenta and $\omega_0 \rightarrow \omega_{\mathbf{k}}$, which means that the Green's functions are momentum dependent (so are all the ϕ s). The Green's Functions are

$$\begin{aligned} G_0^{\text{R}}(\mathbf{k}, t) &= (\epsilon - \omega_{\mathbf{k}} + i0^+)^{-1}, & G_0^{\text{A}}(\mathbf{k}, t) &= (\epsilon - \omega_{\mathbf{k}} - i0^+)^{-1}, \\ G_0^{\text{K}}(\mathbf{k}, t) &= -2\pi i F(\epsilon) \delta(\epsilon - \omega_{\mathbf{k}}) \end{aligned} \quad (10.196)$$

Now introducing interactions

$$H_{\text{int}} = \frac{1}{2} \sum_{qkk'} U_q b_k^\dagger b_{k'}^\dagger b_{k'+q} b_{k-q} \quad (10.197)$$

This is a normal-ordered Hamiltonian, which means that when it's inside the action we can just replace operators with ϕ and $\bar{\phi}$.

$$S_{\text{int}} = -\frac{g}{2} \sum_{qkk'} \oint_C dt \bar{\phi}_{k,t} \bar{\phi}_{k',t} \phi_{k'+q,t} \phi_{k-q,t} \quad (10.198)$$

Splitting the contour into the forward and backward branch then performing the Keldysh rotation gives us

$$S_{\text{int}} = -\frac{g}{2} \int_{-\infty}^{\infty} dt \int d\mathbf{r} [\bar{\phi}^{\text{cl}} \bar{\phi}^{\text{q}} \phi^{\text{cl}} \phi^{\text{cl}} + \bar{\phi}^{\text{cl}} \bar{\phi}^{\text{q}} \phi^{\text{q}} \phi^{\text{q}} + c.c.] \quad (10.199)$$

Hence the interaction introduces the following vertices

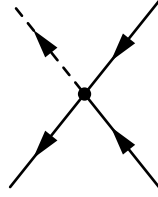


Figure 10.6: Vertex introduced by the $\bar{\phi}^{\text{cl}} \bar{\phi}^{\text{q}} \phi^{\text{cl}} \phi^{\text{cl}}$ term

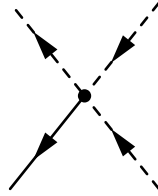


Figure 10.7: Vertex introduced by the $\bar{\phi}^{\text{cl}} \bar{\phi}^{\text{q}} \phi^{\text{q}} \phi^{\text{q}}$ term

The two complex conjugates are just the time reversed processes, so where we swap the directions of all the arrows. These extra terms in the partition function do not contribute – all additional terms are zero.

Keldysh-Dyson Equation

The *full* Green's function is given by

$$\underline{G}^{\alpha\beta}(\mathbf{k}, \mathbf{k}', t, t') = -i \int \mathcal{D}[\phi^{\text{cl}}, \phi^{\text{q}}] \phi^{\alpha}(\mathbf{k}, t) \bar{\phi}^{\beta}(\mathbf{k}', t') e^{iS[\phi^{\text{cl}}, \phi^{\text{q}}]} \quad (10.200)$$

where $S = S_0 + S_{\text{int}}$. If we now expand in terms of S_{int} we can obtain a Dyson Equation:

$$\underline{G} = \underline{G}_0 + \underline{G}_0 \circ \underline{\Sigma} \circ \underline{G} \quad (10.201)$$

where the self-energy has the same causal structure as \underline{G}_0^{-1} , that is

$$\underline{\Sigma} = \begin{pmatrix} 0 & \Sigma^{\text{A}} \\ \Sigma^{\text{R}} & \Sigma^{\text{K}} \end{pmatrix} \quad (10.202)$$

Moving this around we obtain

$$(\underline{G}_0^{-1} - \underline{\Sigma}) \circ \underline{G} = \mathbb{1} \quad (10.203)$$

We can write this out in terms of the components, starting with the retarded and advanced parts

$$(G_0^{-1} - \Sigma^{R(A)}) \circ G^{R(A)} = 1 \quad (10.204)$$

written out

$$\left(i\partial_t + \frac{1}{2m} \nabla_{\mathbf{r}}^2 - V^{\text{cl}}(x) - \Sigma^{R(A)} \right) \circ G_{x,x'}^{R(A)} = \delta^{(4)}(x - x') \quad (10.205)$$

where $x = (\mathbf{r}, t)$. However, there is also a Keldysh component

$$(G_0^{-1} - \Sigma^R) \circ G^K - \Sigma^K \circ G^A = 0 \quad (10.206)$$

using our Ansatz that $G^K = G^R \circ F - F \circ G^A$, after some manipulation we get that

$$[F \circ G_0^{-1}] = \Sigma^K - (\Sigma^R \circ F - F \circ \Sigma^A) \quad (10.207)$$

written out we get the quantum kinetic equation for the distribution matrix, F :

$$-\left[\left(\partial_t + \frac{1}{2m} \nabla_{\mathbf{r}}^2 - V^{\text{cl}}(x) \right) \circ F \right] = \Sigma^K - (\Sigma^R \circ F - F \circ \Sigma^A) \quad (10.208)$$

The left-hand side is the kinetic term – what was referred to as the *streaming term* for the Boltzmann Equation, whereas the right-hand side is the collision integral. In equilibrium the collision term disappears, which means that the Keldysh self-energy obtains the same form as our Ansatz for the Keldysh Green's function.

Let us write out the different components in the commutator

$$[i\partial_t \circ F] = (i\partial_t + i\partial_{t'})F(t, t') \quad (10.209)$$

$$[\nabla_{\mathbf{r}}^2 \circ F] = (\nabla_{\mathbf{r}}^2 - \nabla_{\mathbf{r}'}^2)F(\mathbf{r}, \mathbf{r}') \quad (10.210)$$

$$[V(x) \circ F] = [V(x) - V(x')]F(x, x') \quad (10.211)$$

10.5.7 Wigner-Transformation

Let us consider a function of the form

$$A(x_1, x_2) = A(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) \quad (10.212)$$

Then the Wigner-Transformation is

$$A(x, p) \equiv \int dx' e^{-ipx'} A(x + x'/2, x - x'/2) \quad (10.213)$$

where $x = \frac{1}{2}(x_1 + x_2)$ and $x' = x_1 - x_2$. The inverse transformation would be

$$A(x_1, x_2) = \sum_{\mathbf{p}} e^{ip(x_1 - x_2)} A\left(\frac{1}{2}(x_1 + x_2), p\right) \quad (10.214)$$

We will use this for the quantum kinetic equation, so it is useful to know what the Wigner-Transformation of a convolution is

$$C(x_1, x_2) = \int dx' A(x_1, x')B(x', x_2) \quad (10.215)$$

We find (see Section 10.A at the end of the document)

MOYAL PRODUCT

$$C(x, p) = A(x, p) \exp \left[\frac{i}{2} (\tilde{\partial}_x \tilde{\partial}_p - \tilde{\partial}_p \tilde{\partial}_x) \right] B(x, p) \quad (10.216)$$

remember x and p are four-vectors. Let us suppose that the functions A and B vary in x and p with characteristic scales δx and δp respectively, then it's sufficient to truncate the exponential after the first term, that is

$$\partial_x \partial_p = \nabla_{\mathbf{r}} \nabla_{\mathbf{k}} - \partial_t \partial_\epsilon$$

$$C(x, p) \approx AB + \frac{i}{2} (\partial_x A \partial_p B - \partial_p A \partial_x B) + O(\delta x^2 \delta p^2) \quad (10.217)$$

This is the so-called *Gradient Expansion*, which can be written in terms of Poisson-Brackets:

$$C(x, p) \approx A(x, p) B(x, p) + \frac{i}{2} \{A, B\}_{\text{PB}} \quad (10.218)$$

note that this implies that the commutator becomes

$$[A \oslash B] \xrightarrow{\text{Wigner-Transformation}} \approx i \{A, B\}_{\text{PB}} \quad (10.219)$$

Now, let us use the Gradient Expansion on the kinetic terms from the quantum kinetic equation:

$$\begin{aligned} [i\partial_t \oslash F] &\rightarrow i\partial_t F \\ [-\partial_t^2 \oslash F] &\rightarrow 2i\omega \partial_t F \\ [-\nabla_{\mathbf{r}}^2 \oslash F] &\rightarrow -2i\mathbf{k} \cdot \nabla_{\mathbf{r}} F \\ [\omega_{\mathbf{k}}^2 \oslash F] &\rightarrow -2i\omega_{\mathbf{k}} \mathbf{v}_{\mathbf{k}} \cdot \nabla_{\mathbf{r}} F \end{aligned} \quad (10.220)$$

on the other hand the collision integral becomes

$$\Sigma^K - F(\Sigma^R - \Sigma^A) - i(\partial_x \Re \Sigma^R) \partial_p F + i(\partial_p \Re \Sigma^R) \partial_x F \quad (10.221)$$

now combining it all and moving things around

$$[(1 - \partial_\omega \Re \Sigma^R) \partial_t + (\partial_t \tilde{V}) \partial_\omega + \tilde{\mathbf{v}}_{\mathbf{k}} \cdot \nabla_{\mathbf{r}} - (\nabla_{\mathbf{r}} \tilde{V}) \cdot \nabla_{\mathbf{k}}] F = \mathcal{I}_{\text{coll}}[F] \quad (10.222)$$

where

$$\begin{aligned} \tilde{V} &= V(x) + \Re \Sigma^R(x, p) \\ \tilde{\mathbf{v}}_{\mathbf{k}} &= \nabla_{\mathbf{k}} [\omega_{\mathbf{k}} + \Re \Sigma^R(x, p)] \\ \mathcal{I}_{\text{coll}}[F] &= i\Sigma^K + 2F\Im \Sigma^R \end{aligned} \quad (10.223)$$

Changing variables in F such that $\omega \rightarrow \omega - \omega_{\mathbf{k}} - \tilde{V}$ and using that $\omega = \epsilon_{\mathbf{k}} + \tilde{V} + \Re \Sigma^R(\omega)$, which we solve for ω we obtain

$$[\tilde{Z}^{-1} \partial_t + \tilde{\mathbf{v}}_{\mathbf{k}} \cdot \nabla_{\mathbf{r}} - (\nabla_{\mathbf{r}} \tilde{V}) \cdot \nabla_{\mathbf{k}}] F = \mathcal{I}_{\text{coll}}[F] \quad (10.224)$$

which looks a lot like the semi-classical Boltzmann Equation.

Jens writes "Since F came from $G^K = (G^R - G^A)F + \dots$, having well-defined quasiparticles, i.e. a peaked spectral function"

IO.A MOYAL PRODUCT

We start with

$$C(x_1, x_2) \equiv \int dx_3 A(x_1, x_3) B(x_3, x_2) \quad (10.225)$$

and we are interested in the Wigner-Transformation thereof. The Wigner-Transformation of a function $A(x_1, x_2)$ is:

$$A(x, p) \equiv \int dx' e^{-ipx'} A(x + x'/2, x - x'/2) \quad (10.226)$$

and the inverse is

$$A(x_1, x_2) = \sum_p e^{ip(x_1 - x_2)} A\left[\frac{x_1 + x_2}{2}, p\right] \quad (10.227)$$

Let us plug in our Wigner-Transformation of A and B into the definition of C :

$$C(x_1, x_2) \equiv \int dx_3 \sum_{p_1} e^{ip_1(x_1 - x_3)} A\left[\frac{x_1 + x_3}{2}, p_1\right] \sum_{p_2} e^{ip_2(x_3 - x_2)} B\left[\frac{x_3 + x_2}{2}, p_2\right] \quad (10.228)$$

If we once again define $x_1 = x + x'/2$ and $x_2 = x - x'/2$.

$$C(x + x'/2, x - x'/2) = \int dx_3 \sum_{p_1} e^{ip_1(x + x'/2 - x_3)} A\left[\frac{x + x'/2 + x_3}{2}, p_1\right] \sum_{p_2} e^{ip_2(x_3 - x + x'/2)} B\left[\frac{x_3 + x - x'/2}{2}, p_2\right] \quad (10.229)$$

Thus it must be the case that

$$C(x, p) = \int dx' e^{-ip'x'} \int dx_3 \sum_{p_1} e^{ip_1(x + x'/2 - x_3)} A\left[\frac{x + x'/2 + x_3}{2}, p_1\right] \sum_{p_2} e^{ip_2(x_3 - x + x'/2)} B\left[\frac{x_3 + x - x'/2}{2}, p_2\right] \quad (10.230)$$

changing variables once again

$$x_A \equiv x_3 - x + x'/2, \quad x_B \equiv x_3 - x - x'/2, \quad p_A = p_1 - p, \quad p_B = p_2 - p \quad (10.231)$$

giving us

$$C(x, p) = \iint dx_A dx_B \sum_{p_A, p_B} e^{i(p_B x_A - p_A x_B)} A[x + x_A/2, p + p_A] B[x + x_B/2, p + p_A] \quad (10.232)$$

expanding A and B in terms of p_A and p_B :

$$A(x, p + p_A) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial}{\partial p} A(x, p) p_A^n \quad (10.233)$$

So the integral

$$\sum_{p_A} e^{-ip_A x_B} p_A^n = (i)^n \delta^{(n)}(x_B), \quad \sum_{p_B} e^{ip_B x_A} p_B^n = (-i)^n \delta^{(n)}(x_A) \quad (10.234)$$

However, in order to evaluate the n^{th} derivative of the delta-function we need to differentiate the rest of the integrand w.r.t the delta-function's argument (partial integration). So expanding A gives us a n^{th} derivative of A w.r.t p_A and an n^{th} derivative of B w.r.t x_B (the chain rule gives us factor 2 each time we take an x -derivative). So we can write this as

$$C(x, p) = \left[\sum_{mn} \frac{1}{n!m!} \left(\frac{i}{2} \partial_{p_A} \partial_{x_B} \right)^n \left(-\frac{i}{2} \partial_{p_B} \partial_{x_A} \right)^m A(x_A, p_A) B(x_B, p_B) \right]_{x_A=x_B=x}^{p_A=p_B=p} \quad (10.235)$$

which we can write as

$$C(x, p) = A(x, p) \exp \left[\frac{i}{2} (\tilde{\partial}_p \vec{\partial}_x - \tilde{\partial}_x \vec{\partial}_p) \right] B(x, p) \quad (10.236)$$

CONTENTS

1	Statistical Physics	3
1.1	Multiplicity, Entropy etc.	3
1.2	Boltzmann Distribution & Helmholtz Free Energy	5
1.3	Thermal Radiation	6
1.4	Chemical Potential	8
1.5	Fermi & Bose Gases	12
1.6	Gibbs Free Energy	14
1.7	Phase Transitions	15
1.A	Appendix	17
2	Mathematical Methods in Physics III	19
2.1	Groups	19
2.1.1	Cyclic subgroups	19
2.1.2	Direct product of groups	20
2.1.3	Multiplication tables	20
2.1.4	Homomorphism and Isomorphism	21
2.1.5	Important fact about groups	21
2.1.6	Finite Groups	21
2.1.7	Rules for multiplying cycles:	22
2.1.8	Equivalence Classes	22
2.1.9	Dihedral group D_n	22
2.1.10	The quaternionic group Q	23
2.1.11	The invariant subgroup	23
2.1.12	Derived subgroups	23
2.1.13	Rotations	23
2.1.14	Lie Algebra	24
2.2	Representation Theory & Character Tables	25
2.2.1	Reducible and irreducible representations	26
2.2.2	The Great Orthogonality Theorem:	26
2.2.3	Regular representation	27
2.2.4	Reality	28
2.2.5	Number of square roots	29
2.3	Rotation Groups and Tensors $\mathbf{SO}(3)$	29
2.3.1	Rank-n tensors in $\mathbf{SO}(N)$	31
2.3.2	Dimensions of rank-n tensors, whose indices can take on N values	32
2.3.3	Antisymmetric	32
2.3.4	Symmetric Traceless	32
2.3.5	$\mathbf{SO}(3)$	32
2.3.6	Adjoint representation	32
2.3.7	Ladder operators and $\mathfrak{so}(3)$	33
2.3.8	Multiplying two representations together (still $\mathbf{SO}(3)$)	33
2.3.9	How do we name irreducible representation?	34
2.3.10	But which states?– Clebsch-Gordon decomposition	34
2.4	Special Unitary Groups $\mathbf{SU}(N)$	35
2.4.1	Generators of $\mathbf{SU}(N)$	37

2.4.2	Adjoint representation of $SU(N)$	37
2.5	$SU(3)$	38
2.5.1	Dimensions of $SU(3)$ tensors	38
2.5.2	Multiplying irreducible representations of $SU(3)$	38
2.6	Weights and Roots - $SU(3)$	39
2.6.1	Root vectors	40
2.7	Weights and Roots – General Lie Algebras	42
2.7.1	Generally for $SO(2l)$	43
2.7.2	And now $SO(2l + 1)$	44
2.7.3	The roots of $SU(N)$	44
2.7.4	Symplectic Algebras ($Sp(2l)$)	45
2.8	Classification of Lie Algebras	46
2.8.1	Cartan Subalgebra	46
2.9	Dynkin Diagrams	48
2.9.1	Cartan Matrix	49
2.10	Things no one can remember (Appendix)	49
2.10.1	Integers modulo n — \mathbb{Z}_n	49
2.10.2	Multiplication table	50
2.10.3	Character table Q	50
2.10.4	Further Weight Diagrams	52
2.10.5	Root diagrams for rank-2 Lie Algebras	53
2.10.6	A Few Dynkin Diagrams	54
3	Atomic Physics	55
3.1	Hydrogen Atom	55
3.1.1	Bohr's Model	55
3.1.2	Einstein's A and B coefficients	56
3.2	Interaction of atoms with radiation	57
3.2.1	Oscillating electric field	57
3.2.2	Rotating wave approximation	58
3.2.3	Einstein's B coefficient	58
3.3	Interaction with Monochromatic Radiation	59
3.3.1	Bloch Sphere and Bloch Vector	60
3.3.2	Optical Bloch Equations	61
3.3.3	Optical absorption cross-section	61
3.3.4	Peak radiative broadening	62
3.3.5	Saturation Intensity	62
3.3.6	Power broadening	62
3.4	Transitions	63
3.4.1	π transitions	63
3.4.2	σ transitions	63
3.5	Fine Structure	64
3.5.1	Spin-orbit coupling	64
3.5.2	Relativistic kinetic energy	65
3.5.3	Darwin Term	65
3.5.4	Lamb shift	65
3.6	Helium	65
3.6.1	Excited states	66
3.6.2	Spin states	66
3.7	Alkalis	67
3.7.1	The quantum defect	67

3.7.2	Central-field approximation	67
3.7.3	Self-consistent solutions	68
3.7.4	Fine structure in the alkalis	68
3.8	<i>LS</i> - and <i>jj</i> -Coupling Scheme	69
3.8.1	Spin-orbit interaction in the <i>LS</i> -coupling scheme	70
3.8.2	Selection rules in the <i>LS</i> -coupling scheme:	71
3.8.3	Zeeman effect in the <i>LS</i> -coupling scheme	71
3.9	Hyperfine Structure and Isotope Shift	72
3.9.1	Isotope shift	72
3.9.2	Zeeman effect and hyperfine structure	73
3.9.3	Measurement of the Zeeman effect	74
3.10	Caesium Atomic Clocks	75
3.11	Diatomic Molecules	75
3.11.1	Energy contributions	75
3.11.2	Diatomic electric dipole transitions	77
3.12	Doppler Free Laser Spectroscopy	77
3.12.1	Doppler broadening	77
3.12.2	The crossed-beam method	78
3.12.3	Saturated Absorption Spectroscopy	78
3.12.4	Two-photon spectroscopy	79
3.13	Laser cooling and trapping	79
3.13.1	Scattering Force	79
3.13.2	Slowing an atomic beam	80
3.13.3	Chirp Cooling	80
3.13.4	Optical Molasses Technique	80
3.13.5	Doppler cooling limit	81
3.13.6	Magneto-optical trap	81
3.A	Appendix	82
3.A.1	Summary of atomic units	82
4	Condensed Matter Physics 1	83
4.1	Specific heat of solids	83
4.1.1	Einstein Model of solids	83
4.1.2	Debye Model of solids	84
4.2	Vibrations of a one-dimensional monatomic chain	86
4.2.1	Properties of the one-dimensional chain	87
4.2.2	Quantum Modes: Phonons	88
4.3	Vibrations of a One-Dimensional Diatomic Chain	88
4.3.1	Normal modes of the diatomic chain	89
4.4	Electrons in Metals	90
4.4.1	Drude theory	90
4.4.2	Thermal transport	91
4.4.3	Sommerfeld theory	91
4.5	Chemical Bonds	93
4.5.1	Covalent Bond	93
4.5.2	Van der Waals, Fluctuating Dipole Forces, or Molecular Bonding	94
4.5.3	Metallic Bonding	94
4.5.4	Hydrogen Bonding	94
4.6	Tight Binding Chain	95
4.6.1	1D chain	95
4.6.2	Filling bands	96

4.7	Crystal Structure	97
4.8	The Reciprocal Lattice	99
4.8.1	Fourier transform of the direct lattice	99
4.8.2	Reciprocal Lattice Points as Families of Lattice Planes	100
4.8.3	Brillouin Zones	100
4.9	Wave Scattering by Crystals	101
4.10	Electrons in a Periodic Potential	102
4.10.1	Nearly Free Electron Model	102
4.10.2	Bloch's Theorem	103
4.11	Insulator, Semiconductor or Metal	104
4.11.1	Energy bands in two and three dimensions	104
4.11.2	Failures of the Band-structure picture of metals	105
4.11.3	Optical properties of metals and insulators	105
4.12	Semiconductor Physics	106
4.12.1	Holes	106
4.12.2	Effective Mass of Electrons and Holes	106
4.12.3	Momentum of holes	107
4.12.4	Adding electrons or holes: Doping	107
4.12.5	Statistical Physics of Electrons and Holes	108
4.13	Semiconductor Devices	109
4.13.1	Designing Band Gaps	109
4.13.2	pn unction	109
4.13.3	Transistor	110
5	Condensed Matter Physics 2	113
5.1	Atomic Physics	113
5.2	Isolated Magnetic Moments	115
5.3	Crystal Fields	117
5.4	Interactions	118
5.5	Order and Magnetic Structures	120
5.6	Landau Theory	122
5.7	Mean-Field Theory	123
5.8	Magnetism in Metals	124
5.9	Superconductivity	130
5.10	Integer Quantum Hall Effect	139
6	General Relativity & Cosmology	141
6.1	Special Relativity	141
6.2	Equivalence Principle	142
6.2.1	Gravitational Redshift	143
6.3	General Spacetimes	143
6.3.1	Local Inertial System	145
6.4	Tensors and the Principle of Covariance	145
6.4.1	Covariant Derivative	146
6.4.2	Acceleration in General Relativity	147
6.4.3	Maxwell's Equations in a general spacetime	148
6.5	Riemann Curvature Tensor	148
6.5.1	Symmetries of the Riemann Curvature Tensor	149
6.5.2	Bianchi Identity	150
6.6	Einstein's Equations	150
6.6.1	Conservation of the Energy-Momentum Tensor	151

6.6.2	Einstein's Equations	151
6.7	Schwarzschild Metric	152
6.7.1	Geodesics	153
6.7.2	Deflection of Light	155
6.8	Black Holes	155
6.8.1	Schwarzschild Metric	155
6.8.2	Kerr Black Hole	159
6.8.3	Asymptotically Flat metrics	161
6.8.4	Black Hole uniqueness	161
6.8.5	Cosmic Censorship Hypothesis	161
6.8.6	Black Hole Mechanics	161
6.9	Cosmology	163
6.9.1	Friedrich-Lemaître-Robertson-Walker metric	163
6.9.2	Hubble's Law	164
6.9.3	Friedmann Equations	165
6.9.4	Evolution of the Scale Factor	167
6.10	Linearised Gravity	169
6.10.1	Gravitational Waves	170
7	Advanced Quantum Mechanics	171
7.1	Fundamental Concepts	171
7.1.1	Stern-Gerlach Experiment	171
7.1.2	Bras, Kets and Operators	171
7.1.3	Wave Functions	174
7.2	Quantum Dynamics	175
7.2.1	Heisenberg Picture	176
7.2.2	Harmonic Oscillator	176
7.2.3	Interpretation of the wavefunction	177
7.2.4	Propagators	178
7.2.5	Gauge Transformations in Electromagnetism	179
7.2.6	Aharonov-Bohm Effect	180
7.3	Theory of Angular Momentum	180
7.3.1	Rotations of Spin- $\frac{1}{2}$ systems	181
7.3.2	Eigenvalues and Eigenstates of Angular Momentum	183
7.4	Symmetry in Quantum Mechanics	185
7.4.1	Symmetry	186
7.4.2	Discrete Symmetries	186
7.5	Approximation Methods	190
7.5.1	Perturbation Theory: Non-Degenerate	190
7.5.2	Perturbation Theory: Degenerate	192
7.5.3	Perturbation Theory: Time-Dependent	192
7.5.4	Fermi's Golden Rule	194
7.5.5	Harmonic Perturbation	195
7.5.6	Energy Shift and Decay Width	195
7.6	Scattering Theory	196
7.6.1	Finding the T -Matrix	197
7.6.2	Scattering Amplitude	198
7.6.3	Born Approximation	199
7.7	Identical Particles	199
7.7.1	Symmetrisation Postulate	200
7.7.2	Quantum Fields	201

7.7.3	Quantisation of the Electromagnetic Field	203
7.8	Relativistic Quantum Mechanics	203
7.8.1	Klein-Gordon Equation	203
7.8.2	Dirac Equation	204
8	Elementary Particle Physics	207
8.1	Quantisation of the Electromagnetic Field	207
8.2	Dirac Equation	208
8.2.1	Useful γ -matrix identities	209
8.2.2	Dirac Conjugation	209
8.2.3	Spin & Lorentz Transformations	210
8.2.4	Weyl Equation	211
8.2.5	Parity Transformation	212
8.2.6	Lorentz Algebra	212
8.2.7	Electromagnetic Interactions	213
8.2.8	Solving the Free Dirac Equation, Rest-Frame	213
8.2.9	Spinors; Positive and Negative Modes	214
8.2.10	Charge Conjugation	215
8.3	Perturbation Theory	215
8.3.1	Fermi's Golden Rule & Coulomb Scattering	215
8.3.2	Cross-Section	217
8.3.3	Scattering off an Arbitrary Potential	217
8.3.4	Electron-Muon Scattering	217
8.3.5	Feynman Diagrams	220
8.4	Fermi Theory	221
8.4.1	Mandelstam Variables	222
8.4.2	Gauge Invariance and Gauge Fields	223
8.4.3	Electroweak Feynman Rules	224
8.5	Yang-Mills Theory	225
8.5.1	Non-abelian Field Strength	226
9	Condensed Matter Theory 1	227
9.1	Quantum Fields	227
9.1.1	Collective Modes	228
9.1.2	Thermodynamic Limit	228
9.1.3	The Continuum Limit	229
9.2	Conserved Particles	229
9.2.1	Coleman's Notation	230
9.2.2	Field Operators in Different Bases	230
9.2.3	Field Operators as Creation and Annihilation Operators	231
9.2.4	Interactions	231
9.3	Simple Example of Second Quantisation	232
9.3.1	Jordan-Wigner Transformation	232
9.3.2	Hubbard Model	233
9.3.3	Fluid of non-interacting fermions, thermal equilibrium	234
9.4	Green's Functions	235
9.4.1	Interaction Picture	235
9.4.2	Time-ordering operator	235
9.4.3	Driven harmonic oscillator	236
9.4.4	Wick's theorem	237
9.4.5	Gell-Mann-Low Theorem	238

9.4.6	Useful quantities calculable from Green's functions	238
9.5	Zero-Temperature Feynman Diagrams	238
9.5.1	Linked-Cluster Theorem	240
9.5.2	Feynman Rules in Momentum Space	240
9.5.3	Hartree-Fock energy	240
9.5.4	Exchange correlation	241
9.5.5	Self-energy	242
9.5.6	Response Functions	243
9.6	Non-Zero-Temperature Many-Body Physics	244
9.6.1	Imaginary-Time Green's Functions	245
9.6.2	Matsubara Representation	246
9.6.3	Contour-Integral Method	246
9.6.4	Generating Functional and Wick's Theorem	247
9.6.5	Electron in a Disordered Potential	248
9.6.6	Electron-Phonon Coupling	249
9.7	Fluctuation-Dissipation Theorem and Linear Response Theory	251
9.7.1	Classical Harmonic Oscillator	251
9.7.2	Fluctuations and Dissipation in a Quantum World	251
10	Condensed Matter Theory 2	253
10.1	Path Integrals	253
10.1.1	Coherent States for Bosons	254
10.1.2	Path Integral for the Partition Function: Bosons	254
10.1.3	Fermions: Coherent States and Grassmann mathematics	256
10.1.4	Path Integral for the Partition Function: Fermions	257
10.1.5	The Hubbard-Stratonovich Transformation	257
10.2	Path Integrals and Itinerant Magnetism	258
10.2.1	Saddle Points and the Mean-Field Theory of the Magnetism	260
10.2.2	Quantum Fluctuations in the Magnetisation	261
10.2.3	Lindhard Function	263
10.3	Superconductivity and BCS Theory	263
10.3.1	Bardeen-Pines	263
10.3.2	The Cooper Instability	264
10.3.3	The BCS Hamiltonian	265
10.3.4	Mean-Field Description of the Condensate	266
10.3.5	Physical Pictures of BCS Theory	266
10.3.6	The BCS Ground State	269
10.3.7	Quasiparticle Excitations in BCS Theory	269
10.3.8	Path Integral Formulation	271
10.3.9	Nambu-Gor'kov Green's Function	272
10.3.10	From BCS To Ginzburg-Landau	273
10.3.11	Anderson-Higgs Mechanism in BCS-Theory	274
10.3.12	Are We Really Describing a Superconductor?	276
10.4	Boltzmann Equation	276
10.4.1	Collisions: Classical	277
10.4.2	Boltzmann H -Theorem	277
10.4.3	Conservation Laws	278
10.4.4	Collisions: Quantum	278
10.4.5	Linearisation	279
10.4.6	Relaxation-Time Approximation	279
10.5	Non-Equilibrium Keldysh Field Theory	280

10.5.1	Closed Time Contour	280
10.5.2	Bosons	282
10.5.3	Keldysh Action	285
10.5.4	Feynman Diagrams	285
10.5.5	From Keldysh to Langevin	285
10.5.6	From Keldysh to the Boltzmann Equation	288
10.5.7	Wigner-Transformation	290
10.A	Moyal Product	291