1 Optics

Usually, we think of light as a wave. But often, we can simplify and just think of it as a ray. We can do it when the wavelength is much smaller than the size of the objects the light touches.

Going in the other direction, the more complete picture is electro-magnetic optics. We can go even further and look at light as a quantum phenomenon.

We will start out in the simplest model, Ray optics or also called Geometrical optics.

1.1 Ray Optics

1.1.1 Fermat's Principle

One of the underlying principles of Ray optics is **Fermat's principle**:
Light travels the path whose optical path length is extremal to variations in the path

Intuitively, this means that if we consider a path that is infinitesimally close the the actual path, the optical path lenth does not change.
The path is usually a minimum, but that is not always the case. In a homogeneuous medium, where the speed of light is the same everywhere, this is just saying that light travels in a straight line.
Now consider two media, let's say vacuum and glass and Points A, B in each medium
In vacuum, the speed of light is about $c = 3 \times 10^8 \, \mathrm{m} \, / \, \mathrm{s}$ and in glass, it is about 1.5 times slower than that. We then can define the **refractive index** of glass to be the fraction

$$n_{\mathrm{glass}} := \frac{c}{v_{\mathrm{glass}}} \simeq 1.5$$

This allows us to define the **optical path length** of a path to be the quantity

$$D := n \cdot \ell$$

, where l is the physical length of the path.
If we can parametrize the space of variations from a path with a scalar quantity η then **Fermat's Principle** can be reformulated to say

$$\frac{dD(\eta)}{d\eta}\big|_{\eta=0} = 0$$

So what path does the light take if we take the situation depicted in figure ##? We can easily calculate the total optical path length to be

$$D = n_1\sqrt{d_1^2 + h_1^2} + n_2\sqrt{d_2^2 + h_2^2}$$

Because h, d_1, d_2 are fixed and $h_2 = h - h_1$ we can write it using only the parameter h_1 so in order

to find the extremum of $D(h_1)$ we need

$$
\begin{aligned}
0 &= \frac{dD(h_1)}{dh_1} \\
&= \frac{n_1 h_1}{\sqrt{d_1^2 + h_1^2}} - \frac{n_2 h_2}{\sqrt{d_2^2 + (h - h_2)^2}} \\
&= n_1 \sin\theta_1 - n_2 \sin\theta_2
\end{aligned}
$$

From this we get **Snell's Law:**

$$
\boxed{n_1 \sin\theta_1 = n_2 \sin\theta_2}
$$

1.1.2   Lenses

Imagine a situtation, where a point A is emitting light in all directions and you want to "send" all the light rays towards B.
Normally, Fermat's principle would tell us that there is only one path form A to B that the light would reach B.

Now what a lens would do is to change the optical path lengths for some paths such that more than one of them is extremal. This results in "more Light" reaching B. (See Figure 1.6)

One example of this is the **single surface lens**, which an arced surface between two materials. (For example light and glass).
We can describe the surface of the surface using some function $g(h)$, which allows us the write down the optical path length for the straight line and the point going trough the point $(g(h), h)$.

For the straight line, we simply have $D_{ACB} = n_1 d_1 + n_2 d_2$, whereas for the path $AC'B$ we have

$$
D_{AC'B} = n_1 \sqrt{h^2 + (d_1 + g(h))^2} + n_2 \sqrt{h^2 + (d_2 - g(h))^2}
$$

Using the taylor approximation of $(1 + \varepsilon)^\alpha$ for $|\varepsilon| \ll 1$, we have $(1 + \varepsilon)^\alpha \simeq 1 + \alpha\varepsilon$.
Using the so called *Paraxial approximation*, we are considering only rays that make a small angle with the optical axis (ACB). This means that we consider h and $g(h)$ to be small in comparison to d_1 and d_2.
This allows us to simplify the optical path length $D_{AC'B}$. After dividing out the $(d_i + g(h))$ term and taking the taylor approximation of this term we get

$$
\begin{aligned}
D_{AC'B}(h) &= n_1(d_1 + g(h))\sqrt{1 + \frac{h^2}{(d_1 + g(h))^2}} + n_2(d_2 - g(h))\sqrt{1 + \frac{h^2}{(d_2 - g(h))^2}} \\
&\simeq n_1(d_1 + g(h))\left(1 + \frac{h^2}{2(d_1 + g(h))^2}\right) + n_2(d_2 - g(h))\left(1 + \frac{h^2}{2(d_2 - g(h))^2}\right) \\
&= n_1 d_1 + n_2 d_2 - (n_2 - n_2)g(h) + \frac{h^2}{2}\left(\frac{n_1}{d_1} + \frac{n_2}{d_2}\right)
\end{aligned}
$$

where the errors are of order $\varepsilon^2 \propto h^4$.

After taking the difference between the two pathlengths and setting it to zero, we can get the right curvature of the lens surface.

$$\Delta D = D_{AC'B} - D_{ACB} = \frac{h^2}{2}\left(\frac{n_1}{d_1} + \frac{n_2}{d_2}\right) - (n_2 - n_1)g(h) = 0$$

$$\implies g(h) = \frac{h^2}{2(n_2 - n_1)}\left(\frac{n_1}{d_1} + \frac{n_2}{d_2}\right)$$

which is a parabola.

In practice, creating parabolic surfaces can be hard so we can further approximate the surfac as a sphere with Radius $R \gg h$:

$$g(h) = R - \sqrt{R^2 - h^2} \simeq \frac{h^2}{2R}$$

for $\frac{1}{R} = \frac{1}{n_2 - n_1}\left(\frac{n_1}{d_1} + \frac{n_2}{d_2}\right)$.

Next we may consider a light source living infinitely far away from you. Here, it's emitted light rays are coming in paralel to you and we want to focus all light rays into a single point using a lens.

Using our previous equations, we would have $d_1 = \infty$ and $d_2 = f$ for the distance to the focal point we want to get

$$R = \frac{f(n_2 - n_1)}{n_2}, \quad f = \frac{Rn_2}{n_2 - n_1}, \quad \frac{n_2}{f} = \left(\frac{n_1}{d_1} + \frac{n_2}{d_2}\right) \tag{1}$$

Check the images in the slides for lecture 2.

When the light enters the lens, we can have multiple scenarios. If we get a positive focal distance $f > 0$, we wee that the light gets converges into a single point.

If however we get $f < 0$, we get a *virtual focal point* at distance f *before* the lens.

If we look from the other side of the lens, it would look like the light is coming from the virtual focal point.

If we chose a spherical surface, we see that the light rays don't intersect at exactly the same point. We call this phenomenon the *spherical abberations*

$$\Delta D_{\text{err}} = \varphi(n_1, n_2)h\left(\frac{h}{f}\right)^3$$

Next we consider a lens with two surfaces. The first surface with Radius R_1 will have focal length f'. Here we will look at the rays hitting the second surface as coming from a virtual source from

behind the lens and use the equation 1 with $d_1 = -f$ and we use $n_2 = 1, n_1 = n$. We then get

$$\frac{1}{R_2} = \frac{1}{1-n}\left(\frac{n}{d_1} + \frac{1}{d_2}\right)$$

$$= \frac{1}{1-n}\left(-\frac{n-1}{R_1} + \frac{1}{f}\right)$$

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

This accurately describes our Biconvex lense $R_1 > 0, R_2 < 0 \implies f > 0$
This gives rise to the **lensmaker's equation**

$$\boxed{\frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2}}$$

where we an object A at distance d_1 before the lens and Image of the object at point B at distance d_2 after the lens.
This gives us the following sign conventions

$f > 0$	Lightrays converge
$f < 0$	Lightrays diverge
$d_2 > 0$	Real Image
$d_2 < 0$	virtual image

1.1.3   Ray Tracing

If we have an object which does not lie on the lens axis, we can use the three rules to find out, where the image is.
Rule 1 gives us a straight line through the center of the lens. Using either rule 2 or rule 3, we intersect rule 1 with the line that goes parallel until it hits the lens and go through the focal point. For a convex lens, we use the focal point of the other side for rule 2, and for concave lenses, we use the focal point on the same side as the object or rule 2. For rule 3, we use the other focal point

To better understand how big the image (virtual or real) will be, we introduce the **Magnification factor** $M := \frac{h'}{h} = -\frac{d_2}{d_1}$.
It's absolute value tell us its size and the sign tells us the orientation.

If we are interested in seeing things far away from us, we dont' necessarily want to increase Magnification factor as much as we want to increase the angular size of the object. One way to do this is to use multiple lenses.

Let's say an object with angular size $\Delta\theta$ whose rays travel trough lens 1 and form an image h' at focal point f_1. We can then magnify the image by lens 2 whose focal point lies at the image h'. This would give us

$$h' = f_1 \tan(\Delta\theta) \simeq f_1 \Delta\theta$$

$$\Delta\theta' = \arctan\left(\frac{h'}{f_2}\right) \simeq \frac{f_1}{f_2}\Delta\theta$$

A microscope does something even better: Instead of putting the image h' at the focal point f_2, we can place the image h' between the focal point and the lens to create a virtual image that is bigger, but further away than h'.

Some important things to consider is that the lensmaker's equation assumes that the index of refraction of the surrounding medium is air. If we place a lens in, for example water. The properties of a lens can change!

1.2   Wave Optics

When we were talking about Light as Rays, we made quite a few appproximations and in certain cases, these approximations can lead to minor errors. The next step in accuracy comes, when we look at light as a wave

The starting point for this section is the wave equation, which can be derived from the Maxwell equations. In vacuum, they are

1. $\nabla \cdot \vec{E} = 0$

2. $\nabla \cdot \vec{B} = 0$

3. $\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$

4. $\nabla \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$

After taking the curl we get

$$\nabla \times (\nabla \times \vec{E}) = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E}$$
$$= \nabla \times \left(-\frac{\partial \vec{B}}{\partial t} \right) = -\frac{\partial}{\partial t}(\nabla \times \vec{B}) = \mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}$$

And we end end up with the wave equation

$$\nabla^2 \vec{E} = \mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} = \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t}$$

where $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$ is the speed of light in vacuum.

The wave equation has he **superposition principle**, where if we have two solution, $\vec{E}_1$ and $\vec{E}_2$, then any linear combination $a\vec{E}_1 + b\vec{E}_2$ is also a solution.

Next, consider solutions of the form

$$\vec{E}(\vec{r}, t) = U(\vec{r})e^{i\omega t}$$

where the physical field is given be its real part $\text{Re}(\vec{E})$.

Here there are two possible solutionn. The plane wave equation

$$\vec{E}(\vec{r}, t) = E_0 e^{i(\vec{k}\cdot\vec{r} - \omega t)}$$

and the spherical wave equation

$$\vec{E}(\vec{r}, t) = \frac{A}{r} e^{i(kr - \omega t)A}$$

, where the wave vector $\vec{k}$ is in relation with the wavelength λ in its absolute value: $|\vec{k}| = \frac{2\pi}{\lambda}$

1.2.1   Huygen's Principle

After Fermat's Principle, where light is treated as travelling in straight lines, we will introduce **Huygen's Principle**, from which Fermat's Principle can be derived under some assumptions.

Every Point on the wavefront of a wave acts as a secondary source of hemispherical waves that progate in the forward direction

Now let's test this: We imagine two points $A(x_A, y_A), B(x_B, y_B)$ in space and a light ray travelling from A to B. We then add an aperture centered at C inbetween the points A and B. Using Fermat's principle, we would expect the aperture to "see" the light as it should pass through it and then emit another wave towards B The spatial part of the electromagnetic wave will be

$$U_{AC}(\vec{r}) \propto \frac{1}{|\vec{r}_C - \vec{r}_A|} e^{ik|\vec{r}_C - \vec{r}_A|}$$

$$U_{CB}(\vec{r}) \propto \frac{1}{|\vec{r}_B - \vec{r}_C|} e^{ik|\vec{r}_B - \vec{r}_C|}$$

Now at B, we should see the sum of the two waves, so we should have

$$U_B \propto \int_{-a/2}^{a/2} dx \int_{-a/2}^{a/2} \frac{1}{|\vec{r}_C - \vec{r}_A| \cdot |\vec{r}_B - \vec{r}_C|} e^{ik(|\vec{r}_C - \vec{r}_A| + |\vec{r}_B - \vec{r}_C|)}$$

Since we can assume that the aperture is very small, we have $a \ll r_A, r_B$ and we get

$$|\vec{r}_C - \vec{r}_A| = r_A - \frac{x x_A + y y_A}{r_A} + \frac{x^2 + y^2}{2 r_A}$$

Our next approximation is called the **Fraunhofer approximation**, where we drop the last term $\frac{x^2+y^2}{2r_A}$. We can think of this as saying that the (hemi-)spherical wave will appear flat in a small region. We then get

$$U_B \propto \int_{-a/2}^{a/2} dx \int_{-a/2}^{a/2} dy \frac{e^{ik(r_A + r_B)} e^{-ik(Xx + Yy)}}{\left(r_A - \frac{x x_a + y y_A}{r_A}\right)\left(r_B - \frac{x x_B - y y_B}{r_A}\right)}$$

, for $X = \frac{x_A}{r_A} + \frac{x_B}{r_B}$ and $Y = \frac{y_A}{r_A} + \frac{y_B}{r_B}$

In our third approximation, we will drop the term $\frac{xx_a + yy_A}{r_A}$ and $\frac{xx_B - yy_B}{r_A}$ and we get

$$U_B \propto \frac{e^{-k(r_A + r_B)}}{r_A r_B} \int_{-a/2}^{a/2} e^{ik(Xx + Yy)} dxdy = \frac{4e^{ik(r_A + r_B)}}{r_A r_B} \cdot \text{sinc}\left(\frac{kXa}{2}\right) \cdot \text{sinc}\left(\frac{kYa}{2}\right)$$

where $\text{sinc}(x) = \frac{\sin(x)}{x}$. Further, we get

$$U_B \propto \frac{a^2 e^{ik(r_A + r_B)}}{r_A r_B} \text{sinc}\left(\frac{kXa}{2}\right) \text{sinc}\left(\frac{kYa}{2}\right)$$

The intensity will then be $I = |E(\vec{r}, t)|^2 = |U(\vec{r})|^2$ and we have

$$I_B = I_0 \text{sinc}^2\left(\frac{kXa}{2}\right) \text{sinc}\left(\frac{kYa}{2}\right)$$

so we end up with the characteristic distribution with a large peak at the center and smaller peaks periodically in the distance to the center.

If we take the limit as $X = 0$ (which implies $\frac{x_A}{r_A} = -\frac{x_b}{r_B}$, then we get the delta-distribution, which mimics ray-optics!

If we look at the equation again

$$U_B \propto \int_{-a/2}^{a/2} dx \int_{-a/2}^{a/2} dy e^{ik(|\vec{r}_C - \vec{r}_A| + |\vec{r}_B - \vec{r}_C|)}$$

we can see the optical path length $D \simeq r_A + r_B - Xx - Yy$

If we try to restate Fermat's principle $\frac{dD(\eta)}{d\eta}\big|_{\eta=0} = 0$ we get the condition in two variables

$$\frac{\partial D(x)}{\partial x} = -X = 0, \quad \frac{\partial D(y)}{\partial y} = -Y = 0$$

We see that the intensity is proportional to a^4, which we wouldn't expect with ray-optics as the size of the aperture shouldn't matter.

The conditions to recover ray optics from this is that we have large $\gamma := \frac{a}{\sqrt{\lambda s}}$, for $s = r_A = r_B$ or equivalently $a \gg \lambda$

We also must have, that the distance to the source and the detector is $\geq \lambda$

To generelize, we will define our square aperture in terms of a transmission function $\tau(x, y)$, which in our case was

$$\tau(x, y) = \begin{cases} 1, & \text{if} \quad x \in (-a/2, a/2) \text{ and } y \in (-a/2, a/2), \\ 0, & \text{elsewhere} \end{cases}$$

In general for any transmission function we will have

$$U_B \propto e^{ik(r_A+r_B)} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \tau(x,y) e^{-ik(Xx-Yy)}$$

We saw that if we the intensity is proportional to a^4, so the wider our slit is, the weaker the peak intensity is.
The width of the diffration pattern is also proportional to the wavelength.

So, the **ray-optics limit** occurs when the aperture size is much bigger than the wavelength, which implies that the width of the diffraction pattern becomes very small and we recover Ray optics. Now let's talk about imaging resolution. When we have Fraunhofer Diffraction, we assumed that the wave hitting the aperture were planar. This assumes that the wave is coming from infinitely far away.

In reality, we can mimic a wave coming from infinitely far away by putting lenses between the source and the aperture.

If we put the source at the focal point of a convex lens, the lightwaves will look like a plane and we can then use the Fraunhofer approximation.

Looking at a couple situations where we have a lightsource and a lense that focuses the light into a source, but there is an aperture is blocking some of the light. We can mimic this by imagining two lenses, a diverging and a converging one in series.
The first lens cancels the convergent behaviour just before the aperture and the second lens redos the cancellation. If we assume that the two lenses get infintely close together, we obtain the original path again, all the while the wave hits the aparatus perpendicularly. So even if it doesn't look like Fraunhofer approximation applies, by imagining extra lenses, we can assume it does.

1.2.2   Spatial resolution of telescopes

In this section, we want to observe objects that emit light by letting their light go trough an aperture thats far away and looking at the resulting refraction pattern.
Remember that as the aperture gets smaller, the sinc function intensity peaks get broader and broader.
This was a direct consequence of the effect of the fourier transformation, where if you have a shorter signal, you can make less accurate measurements of the frequency spectrum.

So if imagine two stars A_1, A_2 and a lens perpendicular to the optical path from A_1 to the lens and the angle from the second star being φ.

We will get a refraction pattern which is the sum of the refraction patters of each star, with the angle determining how much the patters are shifted.
But in order to even tell that we have two refraction patterns in the firs place, we must have that the intensities of the two patters are shifted in a way that their peak intensities don't overlap.

This can be stated as the **Rayleigh-criteria** which can be formulated in the equation

$$X_2 := \frac{x_{A_2}}{r_{A_2}} + \frac{x_B}{r_B} = \sin\varphi + \sin\theta$$

,where θ is the angle of the outgoing lightrays from the lens.
The angle that satisfies Rayleigh-critera must have

$$\sin\left(\frac{kaX_2(\theta=0)}{2}\right) = \sin\left(\frac{k\sin\varphi a}{2}\right) = 0$$

And we obtain

$$\frac{ka\sin\varphi}{2} = \pi \implies \sin\varphi = \frac{\lambda}{a}$$

This is what we would expect. In order to tell apart to different objects far away, we need a bigger aperture i.e. a bigger telescope.

Next we consider a plane wave entering a double slit, distance d apart and then hitting a lense on the other side with Lens length L.
The transmission function of the double slit will be

$$\tau(x) \propto \delta(x - \frac{d}{2}) + \delta(x + \frac{d}{2})$$

Using the generalized result from the previous section we have that the Field at the lens will be of the form

$$U(x_B) \propto e^{ik(x_b^2+s^2)} \int_{-\infty}^{\infty} dx\,\tau(x) e^{ikx_b \frac{X}{\sqrt{x_B^2+s^2}}}$$

$$\propto e^{ik(x_B^2+s^2)} \cos\left(\frac{kxd}{2\sqrt{x_B^2+s^2}}\right)$$

This means that if the Lens length is small, (i.e. $L \ll \frac{d}{\lambda s}$), the cosine will be just 1, so it will look like there is only one slit.

In order to find out that there are two slits, we must have that the distance is above the **Abbé limit**:

$$\frac{kd\frac{L}{2}}{2\sqrt{(\frac{L}{2})^2+s^2}} = \pi, \quad d \gg \lambda\frac{\sqrt{s^2+(\frac{L}{2})^2}}{L}$$

Which can be interpreted as asking: How far must two objects be apart, in order for a lens to detect that we have two objects.

If we define the *Numerical aperture* $NA = \frac{\frac{L}{2}}{\sqrt{s^2+(\frac{L}{2})^2}}$, we can restate the formula above as

$$d > \frac{\lambda}{2NA}$$

> **Demo: Abbé limit**
>
> We have a light source pointing through a circular hole and a wire grid, through a lens and an adjustable slit in horizontal direction. We notice that even if we rotate the slit, the image stays the same, as the slit width stays the same.
> However, if we tighted the slit, the vertical lines get washed out, whereas the horizontal lies stay sharp.
> If we rotat the slit again, we see that now the horizontal lines are sharp and the vertical lines are blurred.

1.2.3   Polarisation

We first look at a couple different types of polarisatiion, for a light travelling along the $z - axis$, i.e. $\vec{k} \parallel \vec{z}$

We will first look at **Linear polarisation**, where the field always points at a fixed direction. So the field is oscillating along a straight line. The field can be expressed as

$$\vec{E}(\vec{r}, t) = (E_x \hat{x} + E_y \hat{y}) e^{i(kz + \omega t)}$$
$$= E_0 (\cos\theta \hat{x} + \sin\theta \hat{y}) e^{i(kz - \omega t)}$$

, where $E_0 = \sqrt{E_x^2 + E_y^2}$ and $\tan\theta = \frac{E_y}{E_x}$ and $\hat{\varepsilon} = (\cos\theta \hat{x} + \sin\theta \hat{y})$ is called the **polarisation unit vector**.

In a more general form, we can have **Elliptical polarisation**, where the field will be of the form

$$\vec{E}(\vec{r}, t) = \left(E_x \hat{x} + E_y e^{i\varphi} \hat{y} \right) e^{i(kz - \omega t)}.$$

For $\varphi = 0$, we recover linear polarisation, and for $\varphi = \pm\frac{\pi}{2}$ we call that right(left) **circular polarisation**.
Warning, the right-left convention is not universally agreed upon, but in this convention, if we place our right thumb along $\hat{k}$, the other fingers follow the polarisation.

1.2.4   Birefringence

Recall from the previous sections, that we called denoted the Index of refraction as n. This had the effect that the effective wave-length changed depending on the medium or the wave vector.

$$\lambda_0 \to \frac{\lambda_0}{n}, \quad k_0 \to \frac{n}{k_0}$$

If we had a material, where the index of refraction differences in multiple directions, i.e. $n_x \neq n_y$ we get the **birefringence** effect:

$$\vec{E}(\vec{r}, t) = \left[E_x \hat{x} e^{ik_x z} + E_y \hat{y} e^{i(k_y z + \varphi)} \right] e^{-i\omega t}$$

where we call the axis with higher index of refraction the *slow axis.*
If the light travels through a such a material with Length L, then the *accumulated phase difference* will be equal to $(n_x - n_y)k_0L$.

For the special case where $(n_x - n_y)k_0L = \pi$ we call that a *half-wave plate.* And if it equals $\frac{\pi}{2}$, we call that a *quarter-wave plate.*

Next we will consider light reflection/transmission through a change in material. In the case where there are no free charges or currents we have

$$\nabla \cdot \vec{D} = 0, \quad \nabla \times \vec{E} = -\frac{\partial B}{\partial t}$$
$$\nabla \cdot \vec{B} = 0, \quad \nabla \times \vec{H} = \frac{\partial D}{\partial t}$$

For a linear material, we will have

$$\vec{B} = \mu\mu_R\vec{H}, \quad \vec{D} = \varepsilon_0\varepsilon_r\vec{E}$$

And for non-magnetic materials we have $\mu_r = 1$, so

$$\nabla^2 \vec{E} = \varepsilon_r\varepsilon_0\mu_0\frac{\partial^2 \vec{E}}{\partial t^2}$$

which changes the speed of light

$$v = \frac{1}{\sqrt{\varepsilon_r\varepsilon_0\mu_0}} = \frac{c}{n}$$

where $n = \sqrt{\frac{\varepsilon_r}{\mu_r}}$ is the index of refraction.

For plane-waves we get

$$\vec{k} \times \vec{E} = \omega\vec{B} \implies B = \frac{k}{\omega}E = \frac{nk_0}{\omega}E = \frac{n}{c}E$$

In the integral form over the closed surface S of a cylinder with area A and height d we have

$$\int_S \vec{B} \cdot \hat{n}dA = 0$$

As we reduce the height of the surcace: we get that

$$d \to 0: \quad AB_{\perp 1} - AB_{\perp 2} = 0 \implies B_{\perp 1} = B_{\perp 2}$$

Similarly, we have $D_{\perp 1} = D_{\perp 2}$. In the Integral version of the fourth macroscopic maxwell equation, we have

$$\int_C \vec{H} \cdot d\vec{l} = \int_S \frac{\partial D}{\partial t} \cdot \hat{n}dA$$

, where C is the boundary of the surface and S is its area with height d and Length L. So as the height goes to zero, only the parallel components remain and we get

$$d \to 0: LH_{\|1} - LH_{\|2} \implies H_{\|1} = H_{\|2}$$

And similarly, $E_{\|1} = E_{\|2}$.
(Figure 1.40, 1.41) Steve

1.2.5   Fresnel equations

So now we can find out the amplitudes of the transmitting and reflecting waves bouncing off an interface. In the case of the p-polariszed light (E-ield is pointing in the plane of incidence) we have

$$E_i \cos \theta_i - E_r \theta_i = E_t \cos \theta_t$$

where E_i is the ampluditde of the *incoming*, E_r of the *reflected* and E_t of the *transmitted* field. From the relation

$$\vec{H} = \frac{\vec{B}}{\mu_0} = \frac{n}{\mu_0 \omega} \vec{k} \times \vec{E}_0$$

and the boundary condition, we obtain

$$n_1 E_i + n_1 E_r = n_2 E_t$$

Combining the newly found equations, we obtain the **Fresnel equations for p-polarisation**:

$$t_p := \frac{E_t}{E_i} = \frac{2 n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t}$$

$$r_p := \frac{E_r}{E_i} = \frac{n_2 \cos \theta_i - n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t}$$

From Figure 1.37 Steve, we see that the x component of the reflecting E field gets a inus sign, For the s-polarisation, where the Electric field is pointing parallel to the surface (which we will derive as homework exercise) we have

$$t_s := \frac{E_t}{E_i} = \frac{2 n_1 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_t}$$

$$r_s := \frac{E_r}{E_i} = \frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t}$$

where the indices of n_1 and n_2 are switched compared to the p-polarisation. Similar as in the p-polarisation case, we see that the x-component of the reflected B-field gets the minus sign.

From writing

$$A_i e^{i(\vec{k}_i \cdot \vec{r} - \omega_i t)} + A_R e^{i(\vec{k}_r \cdot \vec{r} - \omega_r t)} = A_t e^{i(\vec{k}_t \cdot \vec{r} - \omega_t t)}$$

we obtain Snell's law:

$$n_1 k_0 \sin \theta_i = n_1 k_0 \sin \theta_r = n_2 k_0 \sin \theta_t$$
$$\implies \theta_i = \theta_r \quad \text{and} \quad n_q \sin \theta_1 = n_2 \sin \theta_t$$

If we want to actually solve some problems, we can use the following: If we have any arbitrary polarisation, it can be written as a superposition of p and s polarizations. And in order to relate θ_i and θ_t, we can apply Snell's law.

1.2.6 Brewster's Angle

Starting from the Fresnel we can use some "trigonometric tricks" to get

$$
\begin{aligned}
r_p &= \frac{n_2 \cos \theta_i - n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t} \\
&= \frac{\sin \theta_1 \cos \theta_i - \sin \theta_t \cos \theta_t}{\sin \theta_i + \sin \theta_t \cos \theta_t} \\
&= \frac{\sin(2\theta_i) - \sin(2\theta_t)}{\sin(2\theta_i) + \sin(2\theta_t)} \\
&= \frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)}
\end{aligned}
$$

Notice that the denominator diverges for $\theta_1 + \theta_t = \frac{\pi}{2}$. In that case we have $r_p = 0$. The special angle θ_i that causes this is called **Brewster's angle** θ_B. Using Snells law again we have

$$
\theta_B = \frac{\pi}{2} - \theta_t \implies \cos \theta_B = \sin(\theta_t) = \frac{n_1}{n_2} \sin \theta_B
$$

$$
\implies \theta_B = \arctan(n_2/n_1)
$$

From Figure 1.42 Steve, we can see that for n_2 : index of refractin of Water, the Brewster angle is at abou 55°, where p-polarised Light is not reflected. We also see that the reflected amplitude is always higher for s polarisation than for p-polarisation.

> **Demo:**
>
> We have a laser going through two linear polarizers we can rotate to change the relative angles of the polarizers. The light is then pointed to a photodetecter and the intensity is plotted on a screen.
>
> If both are aligned in the same angle, we see that the intensity is high. As we rotate one polarizer, the intensity decreases until we hit the 90 dgree angle, at which we measure no light. We also see that for angles bigger than 90 degrees, we obtain a periodic pattern, with peak intensity reached every 180 degrees.

1.3 Spectroscopy

Electro-magnetic waves can come in a huge spectrum of wavelengths. When we have a mixture of light, we might be interested in finding out what wavelengths are present.
Prisms are able to separate some frequencies, but they aren't really good to make accurate measurements, as they the index of refraction doesn't differ alot for waves, whose wavelengths only differ by a small amount. We could work the difference in angles using Snell's Law.

A more effective way to seperate by wavelength is using **Gratings**.
Consider an array of small mirrors with centers distance d apart with some gaps inbetween. (Steve Fig. 1.44)
Now if we were assuming that there would be no gaps, we could argue with symmetry and find out

that the angle of incoming and outgoing light is the same. But because we have gaps here, we can't do that.

When looking at the light reflected from neighboring mirrors, we can calculate the their phase difference as

$$\varphi(\lambda) = k \cdot (d \sin \theta_i - d \sin \theta_j), \quad \text{where} \quad k = \frac{2\pi}{\lambda}$$

Now look at the contribution of all N strips and we can see that the Intensitiy will be

$$U \propto \sum_{n=0}^{N-1} e^{i\varphi n} = \frac{1 - e^{iN\varphi}}{1 - e^{i\varphi}}$$

$$= \frac{e^{i\varphi(\frac{N}{2})} - e^{i\varphi\frac{N}{2}}}{e^{i\frac{\varphi}{2}}(e^{-\frac{\varphi}{2}} - e^{i\frac{\varphi}{2}})} = e^{i(N-1)\frac{\varphi}{2}}\left(\frac{\sin(N\frac{\varphi}{2})}{\sin(\frac{\varphi}{2})}\right)$$

So the Intensitiy will have

$$I = |U|^2 \propto \frac{\sin^2\left(N\frac{\varphi}{2}\right)}{\sin^2\left(\frac{\varphi}{2}\right)}$$

which will have its maxima at

$$\frac{\varphi}{2} = m\pi \quad \text{for} \quad m \in \mathbb{Z}$$

At these points, the Intensity will be

$$\lim_{\varepsilon \to 0} \frac{\sin^2\left(N\frac{\varphi}{2} + \varepsilon\right)}{\sin^2\left(\frac{\varphi}{2} + \frac{\varepsilon}{N}\right)} = N^2$$

See (Steve Fig. 1.47)

$$\sin \theta_i - \sin \theta_j = \frac{2\pi m}{dk} = \frac{m\lambda}{d}$$

Now we can ask when we will be able to resolve the difference between two wavelenegths. Looking at the (Fig. 1.47), we can frame the condition (similar to the Rayleigh criterium) as follows:
The conditions for the peak of wavelength λ_1 sitting at the zero of wavelength λ_2 will be

$$\frac{\varphi(\lambda_1)}{2} = m\pi, \quad \text{and} \quad \frac{\varphi(\lambda_2)}{2}N = m\pi N + \pi$$
$$\implies ND(\sin \theta_i - \sin \theta_j)(k_2 - k_1) = 2\pi$$

So the smallest resolvable frequency difference wil be

$$\boxed{\begin{aligned} \Delta\omega_{\min} &= (k_2 - k_1)c \\ &= \frac{2\pi c}{ND(\sin \theta_i - \sin \theta_j)} > \frac{2\pi c}{2Nd} \end{aligned}}$$

So the more Mirrors we have the tighter we can resolve different frequencies.

We have observed that the resolution depends on the total *path-difference*. The next tool will try to maximize this value

The **Michelson Interferometer** will have the setup (Steve Fig. 1.49)
There the path difference will be $2x = 2(d_2 - d_1)$ which can be adjusted however we want. When looking at the field, we have

$$U = \frac{U_0}{2}(1 + 2^{2ikx}) \implies I = |U|^2 = \frac{|U_0|^2}{4}(2 + 2\cos(2kx)) = \frac{I_0}{2}\left(1 + \cos\left(\frac{2\omega x}{c}\right)\right)$$

We see that the shorter the wavelength is, the more sensitive the Interferometer in x. It is in fact so sensitive, that it can be used to detect gravitational waves at LIGO.

In order to talk about the *distribution* of frequences inn a light bundle, we can define the **spectral density** $S(\omega)$ of a bundle of light by measuring the added Intensity of all frequences between ω and $\omega + d\omega$.
So in general we will have

$$I(x) = \frac{I_0}{2}\int_0^\infty S(w)\left(1 + \cos\left(\frac{2\omega x}{c}\right)\right) d\omega$$
$$= \frac{I_0}{2} + \frac{I_0}{2}\int_0^\infty S(\omega)\cos\left(\frac{2\omega x}{c}\right) d\omega$$

We can invert this relation using the inverse fourier/cosine transform to get

$$S(\omega) = \int_0^\infty (2I(x) - I_0)\cos\left(\frac{2\omega x}{c}\right) dx$$

Which corresponds to measuring the intensity as we move our mirror further and further away. In reality, we obviously can't move our mirro infitely far away, so we can only take the integral from 0 to some x_{max}

This means that depending on how much "room" we have for our Intereferometer, we can resolve frequencies up to a certain maximam, with the theoretical limit being infinite precision.

In the ideal case with a monochromatic source with some frequency ω_0, we should get the dirac delta function $S(\omega) = \delta(\omega - \omega_0)$, but in reality we would get

$$S(w) = \int_0^{x_{\text{max}}} \cos\left(\frac{2\omega_0 x}{c}\right)\cos\left(\frac{2\omega x}{c}\right) dx$$
$$\propto \text{sinc}\left(\frac{(\omega - \omega_0)x_{\text{max}}}{c}\right) + \text{sinc}\left(\frac{(\omega + \omega_0)x_{\text{max}}}{c}\right)$$

where the second term is negligible around $\omega \sim \omega_0$. So we will obviously have a peak at $\omega = \omega_0$ with the closst zeros at

$$\Delta\omega = \omega - \omega_0 = \frac{\pi c}{x_{\text{max}}} \quad \text{or} \quad \frac{\Delta\omega}{\omega_0} = \frac{\lambda}{2x_{\text{max}}}$$

Now here the path difference is "only" proportional to twice the physical length x. So how do we get even more path difference?

The **Fabry-Perot etalon** is a device (Steve Fig. 1.50) where we have two cavity mirrors placed parallel to each other. Here the light will keep bouncing between these mirrors. So what do these Amplitudes look like? Every boucen, some part of of it will be transmitted and "lost" and the other part wil be reflected.

So every round trip we will accumulate a nice amount of phase differece at the cost of losing some intensity.

If we write A_0 for the incoming Amplitude, A_1 for the light that gets transmitted in the straight line, and A_2 for the amplitude if it makes one round trip etc, we will have

$$A_1 = A_0 t^2 e^{ikd} \quad , \quad A_2 = A_0 t^2 e^{ikd} r^2 e^{i\varphi}$$
$$A_3 = A_0 t^2 e^{ikd} r^4 e^{2i\varphi}, \quad A_l = A_0 t^2 e^{idk} r^{2(l-1)} e^{i(l-1)\varphi}$$

where the phase φ is a function of the frequency

$$\varphi = 2kd + \varphi_r = \frac{2\omega d}{c} + 2\varphi_r$$

If we set A to be the sum of all out-going amplitudes on the right side, we get

$$A := \sum_{l=1}^{\infty} A_l = A_0 t^2 e^{idk} \sum_{n=0}^{\infty} r^{2n} e^{in\varphi}$$
$$= A_0 t^2 e^{idk} \frac{1}{1 - r^2 e^{i\varphi}}$$

So comparing the Intensities we will have

$$\frac{I_{\text{out}}}{I_{\text{in}}} =: \eta = \frac{|A|^2}{|A_0|^2} = \frac{t^2}{1 + r^4 - 2r^2 \cos\varphi}$$

If we write $T := t^2$ and $R := r^2$ and assume that $t^2 + r^2 = 1$ we have

$$\eta = \frac{(1-R)^2}{(1-R)^2 + 2R(1 - \cos\varphi)} = \frac{1}{1 + \frac{4R}{(1-R)^2} \sin^2\left(\frac{\varphi}{2}\right)}$$

So if we have a higher reflection index R, it means that our light can reflect more often (probabilitstically) and therefore we get a more sensitive result in ω.

Demo: Diffration grating

In this setup, we shine some light through some gratings. If we use some with 80 slits per centimeter, we just an interference patters
At 250 slits per centimeter, we see that the first maxima are still quite white but we can see that there are multiple colours present in the smaller maxima.
At 500 slites per centimeter, the colours are visibly separated from the second maxima on.

Also note that the Michelson Interferometer is sensitive to the index of refraction. With this, we can detect turbulences in the temperature oflight in our optical path.

2   Statistical Physics

As we did quite a bit in Optical physics, we tried to simplify complex problems until we could solve then. Statistical physics is all about simplification. If we have a gas, we know enough physics to describe each molecule individually, but we can't fully describe the system as a whole.
So to make useful statements about the entire system, we ry to extrapolate information from the microscopic system to the macroscopic one.

> **Definition**
>
> A **microstate** is a configuration of all the microscopic properties. For example one might describe position, momentum or velocity of individual molecules but we can never fully the microstate of a system.
> A **macrostate** is a particular set of macroscopic properties. For example the temperature, the pressure, volume of a gas.
> It is clear that a microstate fully determines the macrostate but not the other way around and this property *must be satisfied* to to pair them.

For example, if we have four coins which we can flip and tell apart, a microstate would be the configuration of the individual coins, i.e. $HTTH$ or $TTTH$. A macrostate could consist of a variable that indicates the total number of heads.
In the example above, one can consider the probability that our variable takes on the value 3 for any given microstate. We then can find out the **discrete probability distribution** for the variable over the microstates.
Once we know the probability distribution of a variable z we can calculate the the discrete average value to be the weighted sum over all macrostates:

$$< x >:= \sum_{x\text{microstate}} P(x) \cdot x$$

However, when we have an infinite number of microstates, we can define a variable that can take on any value in an interval. For example if x denotes the position of a particle in a 1 dimensional space, we might ask what the probabilty is for the particle to be inbetween the range x_0 and $x_0 + dx$. Then we have

$$\int_{-\infty}^{\infty} p(x)dx = 1, \quad < x > \int_{-\infty}^{\infty} p(x)xdx$$

where we call $< x >$ the **mean value** of the variable x. We can also define the mean value of a function $f(x)$ to be

$$< f(x) >= \int_{-\infty}^{\infty} p(x)f(x)dx$$

With many systems, the macroscopic variables we define tend to stabilize afte some time. We call this **Thermodynamic equilibrium**, which satisfies the following postulates:
The **fundamental postulate** states that for a closed system with a fixed set of global constraints, every microstate that satisfies these constraints have equal probability of begin occupied.

When we talk about a **closed system** we mean that there can be no exchange of matter, but energy might be exchanged with the outside.

The second postulate states the macrostate occupied in the thermodynamic equilibrium is the one with the largest number of microstates.

In our coin-toss example we keep tossing one coin every second and we ask, what is the probability that we end up with the mirostate $TTTT$ at n seconds?

$$P_n(4) = P(4|3)P_{n-1}(3) + P(4|4)P_{n-1}(4)$$

$$P(4|3) = \frac{1}{4} \cdot \frac{1}{2}, \quad P(4|4) = \frac{1}{2}$$

Where $P(a|b)$ denotes the propability that our variable x takes on the variable a *under the condition* that we have 3 tails.
In the equilibrium we have $P_n(z) = P_{n-1}(z)$. From this we find that

$$P_\infty(4) = \frac{\frac{1}{8}P_\infty(3)}{1 - \frac{1}{2}} = \frac{1}{4}P_\infty(3)$$

Similarly we get

$$P_\infty(3) = \frac{2}{3}P_\infty(2), \quad P_\infty(1) = \frac{2}{3}P_\infty(2), \quad P_\infty(0) = \frac{1}{4}P_\infty(1)$$

From the condition that all the probabilites must add up to zero we then get

$$P_\infty(z) = \begin{cases} \frac{1}{16} & \text{for } z = 0, 4 \\ \frac{4}{16} & \text{for } z = 1, 3 \\ \frac{6}{16} & \text{for } z = 2 \end{cases}$$

For the general problem with N coins we the $< x >= \frac{N}{2}$, where the width of the peak is given by the standard deviation

$$\Delta z = \sqrt{< z^2 > - < z >^2} = \frac{N}{2}$$

Where the variance is $\Delta z^2 = \frac{N}{4}$ and the fractional width is $\frac{\Delta z}{<z>} \propto \frac{1}{\sqrt{N}}$

When we want to measure a property of a system, we often do this by making it interact with another system and see how the second system reacts to it.
If we write N_i, U_i for the number of particles and the Energy of each system, we see that $U_0 := U_1 + U_2$ is conserved.

Let $\Omega_1(N_1, U_1)$ denote the number of microstates that are consistent with N_1 and U_1. Similarly Ω_2. Then we have

$$\Omega(N_1, N_2, U_1, U_2) = \Omega_1(N_1, U_1) \cdot \Omega_2(N_2, U_2)$$

be the total number of microstates for the entire system. If we pick U_1 as a variable, we have

$$\left(\frac{\partial \Omega_1}{\partial U_1}\right)_{N_1} \Omega_2 + \left(\frac{\partial \Omega_2}{\partial U_2}\right)_{N_2} \left(\frac{\partial U_2}{\partial U_1}\right)_{N_2} \Omega_1 = 0$$

Because U_0 is constant, $\frac{\partial U_2}{\partial U_1} = -1$ so

$$\frac{1}{\Omega_1} \left(\frac{\partial \Omega_1}{\partial U_1}\right)_{N_1} = \frac{1}{\Omega_2} \left(\frac{\partial \Omega_2}{\partial U_2}\right)_{N-2}$$

$$\implies \left(\frac{\partial (\ln(\Omega_1))}{\partial U_2}\right)_{N_1} = \left(\frac{\partial (\ln(\Omega_2))}{\partial U_2}\right)_{N_2}$$

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

for $\sigma = \ln(\Omega)$). This gives us a relation between the two system when thermodynamic equilibrium is reached. Note that the left and right hand side only depend on the variables associated to only one system. This means that for any two systems that come into contact with eachother, the quantity $\left(\frac{\partial \sigma}{\partial U}\right)$ is the same for both after the thermodynamic equilibrium! This property also defines the *Temperature*, since after enough time, both systems have the same temeprature. We use this to give a proper definition of Temperature

$$\frac{1}{T} = k_B \left(\frac{\partial \sigma}{\partial U}\right)_{N,V}$$

where $k_B = 1.38 \cdot 10^{-23} m^2 kg/s^2 K$ is the **Boltzmann constant**. Since temperature and the Boltzmann constant are often paired up, we define β to be

$$\beta := \frac{1}{k_B T} = \left(\frac{\partial \sigma}{\partial U}\right)_{N,V}$$

We define the **Entropy** S in the Boltzmann equation

$$S = k_B \sigma = k_B \ln(\Omega)$$

2.1   Laws of thermodynamics

Zeroeth Law: Two systems in thermal equilibrium have the same temperature.

First Law: Heat is a form of energy and the change in Energy ΔU is the sum of applied heat $Q^{\checkmark}$ and the work done on the system $\delta W^{\checkmark}$. A reversible process is when all systems involved remain in thermal equilibrium. For a reversible process we have the relationship between Heat, Temperature and Entropy

$$\delta Q = T \ dS$$

Second Law: The Entropy of an insolated system never decreases.

2.2   Boltzman Factor and the partition function

THe first question we will ask is that if we have a system with Temperature T, what is the probability for it to be in a microstate with energy ε.

We do this by attaching the system with a reservoir with a much larger total energy. Such that the total energy $U_0 \gg \varepsilon$ is much larger (Steve Fig 3.7). If we assume that they are isolated from the outside, but heat can flow freely between them, then the total energy is fixed.
Then the number of microstates in which the reservoir has Energy U_b, equals the number of microstates in which the System has the rest of the energy.

$$\Omega(U_b) = \Omega(U_0 - \varepsilon)$$

This means that we can express the probability of finding the system in microstate ε is.

$$P(\varepsilon) = \Omega(U_0 - \varepsilon)/z$$

Expressing this in terms of $\sigma_R(U_R) = \ln \Omega_R(U_R)$ we have that

$$\frac{P(\varepsilon_1)}{P(\varepsilon_2)} = \frac{\Omega_R(U_0 - \varepsilon_1}{\Omega_R(U_0 - \varepsilon_2)} = e^{\sigma_R(U_0 - \varepsilon_1) - \sigma_R(U_0 - \varepsilon_2)}$$

Using $\varepsilon_1, \varepsilon_2 \ll U_0$ we get

$$\frac{P(\varepsilon_1)}{P(\varepsilon_2)} \simeq e^{-(\frac{\partial \sigma_R}{\partial U})_{N_R}(\varepsilon_1 - \varepsilon_2)} \frac{e^{-\beta \varepsilon_1}}{e^{-\beta \varepsilon_2}}$$

where $\beta = (\frac{\partial \sigma}{\partial U}) = \frac{1}{k_B T}$. By defining the **Boltzmann factor** to be $e^{-\beta \varepsilon}$, we get the **partition function** Z, which sums over all possible microstates s of the system.

$$Z = \sum_s e^{-\beta \varepsilon_s}$$

so the probability of finding the system in the microstate with energy ε we get

$$P(\varepsilon) = \frac{e^{-\beta \varepsilon}}{Z}$$

Note that we have the folllowing property:

$$\frac{\partial \ln Z}{\partial x} = \frac{1}{Z}\frac{\partial Z}{\partial x} = \frac{1}{z}\sum_s \frac{\partial(\beta\varepsilon_s)}{\partial x}e^{-\beta\varepsilon_s} = \left\langle \frac{\partial(-\beta\varepsilon_s)}{\partial x} \right\rangle$$

We can make use of this to calculate the Energy U:

$$U = \langle \varepsilon_s \rangle - \frac{\partial \ln(Z)}{\partial \beta}$$

The second question we will be asking is the following:
If we have N atoms with total energy U, and each atom can be in one f a number of states, labeled by i, where state i has energy ε_i and $E \gg \varepsilon_i$.
In equilibrium, what is the probability p_i of a particular atom is in the state i?

If we write n_i for the number of atoms in state i, then $\{n_i\}$ defines a macrostate and we am maximaize Ω over n_i to find the equilibrium macrostate.

If we have N atoms, how many ways are there to distribute them to get $\{n_i\}$? It is clear that

$$\Omega = \frac{N}{n_1!(N-N_1)!} \cdot \frac{(N-n_1)!}{n_2!(N-n_1-n_2)!} \cdots \frac{n_k!}{n_k!0!} = \frac{N!}{n_1!n_2!\ldots n_k!}$$

so instead of maximizing Ω we can instead maximize $\ln(\Omega) = \ln(N!) - \sum_i \ln(n_i!)$ subject to the constraints

$$\sum_i n_i = N \quad \sum_i \varepsilon_i n_i = E$$

Using the method of Lagrange multipliers α and β the maximization is found when

$$0 \overset{!}{=} \frac{\partial}{\partial n_i} \ln(n_i!) + \alpha + \beta\varepsilon_i$$

If we assume that n_i is large, using Stirlings approximation we have

$$\ln(x!) \simeq x\ln x - x$$

So for $\ln(n_i) + \alpha + \beta\varepsilon_i$ we have

$$n_i = e^{-\alpha}e^{-\beta\varepsilon_i} \implies p_i = \frac{n_i}{N} = \frac{e^{-\alpha}e^{-\beta\varepsilon_i}}{\sum_i e^{-\alpha}e^{-\beta\varepsilon_i}}$$

If we identify $\beta = \frac{1}{k_B T}$ this describes the situation where our system is a single atom and the reservoir is a liquid or a gas.

2.3   Ideal Gas

In an ideal gas, the energy of a single paticle is just its kintic energy $\varepsilon = \frac{p^2}{2m}$. So for the partitiion function Z we cab ask: what is the probability that a particle has a position between q and $q + dq$ and a momentum between p and $p + dq$?.

In 1 dimension, it will $P(q,p)dqdp$ which is just the scaled area spanned by dq and dp.

In 3 dimensions, we have six dimensions:

$$P(\vec{q}, \vec{p})d^3qd^3p = \frac{1}{A}e^{-\beta\varepsilon}d^3qd^p$$

where

$$A = \int\int e^{-\beta\varepsilon_i}d^3qd^3p = \int d^3q \int e^{-\beta p^2/2m}d^3p$$

We introduce the **Planck's constant** h, which has position $\cdot$ momentum as units.

For a single particle, the partition function will be of the form

$$Z_{\mathrm{sp}} = \frac{A}{h^3} = \frac{V}{h^3}\int_0^\infty 4\pi p^2 e^{-\beta p^2/2m}dp$$

And by using $\int_0^\infty x^2 e^{-ax^2}dx = \frac{\sqrt{\pi}}{4a^{3/2}}$ we find that

$$Z_{\mathrm{sp}} = \frac{v}{h^3}\left(\frac{2\pi m}{\beta}\right)^{3/2}$$

Now if the particles are distinguishable (but identical), we get the many-particle partition function, which can be expressed in terms of the single particle partition function Z_{sp}

$$\begin{aligned}
Z_{\mathrm{D}} &= \sum_{s_1}\sum_{s_2}\ldots\sum_{s_N}e^{-\beta(\varepsilon_{s_1}+\varepsilon_{s_2}+\ldots+\varepsilon_{s_n})} \\
&= \left(\sum_{s_1}e^{-\beta\varepsilon_{s_1}}\right)\left(\sum_{s_2}e^{-\beta\varepsilon_{s_2}}\right)\ldots\left(\sum_{s_N}e^{-\beta\varepsilon_{s_N}}\right) \\
&= Z_{\mathrm{sp}}^N
\end{aligned}$$

where s_i is the state of particle i

2.3.1   Indistinguishable particles

In principle, if we could keep track of all particles in a gas an therefore distinguish them. But since the information is often very limited, we won't be able to do that. (As we will see later, quantum mechanics also makes it impossible to track everything)

Recall that for distinguishable particles, we could write

$$Z_D = Z_{\mathrm{sp}}^N, \quad \text{where} \quad Z_{\mathrm{sp}} = \frac{V}{h^3}\left(\frac{2\pi m}{\beta}\right)^{\frac{3}{2}}$$

which does not hold anymore for indistinguishable particles! For example for 2 distinguishable particles we have

$$
\begin{aligned}
Z_D &= \left(\sum_{s_1} e^{-\beta \varepsilon_{s_1}} \right) \left(\sum_{s_2} e^{-\beta \varepsilon_{s_2}} \right) \\
&= \sum_s e^{-2\beta \varepsilon_s} + \sum_{s_1} \sum_{s_2 \neq s_1} e^{-\beta(\varepsilon_{s_1} + \varepsilon_{s_2})}
\end{aligned}
$$

With the first term corresponding to the cases where both particles are in the same states, and the second term summing over all states, in which the two particles occupy two different states. But since we can't tell the particles apart, we are actually counting the states twice in the above example.

So for N particles, we need to divide out the multiples of the states so we get

$$
Z_I = \sum_s e^{-N\beta \varepsilon_s} + \ldots + \frac{1}{N!} \sum_{s_1} \cdots \sum_{s_N} e^{-\beta(\varepsilon_{s_1} + \ldots + \varepsilon_{s_N})}
$$

where the middle terms describes situations, where two ore more particles occupy the same state s_i. But in the **Classical regime**, we can assume that if we have so many possible states for a single particle to be in, the probabilty that two particles are in the same state is so low, that we can leave them out in our sum, so the only term that remains (if we have enough states) is the last term:

$$
Z_I = \frac{1}{N!} \sum_{s_1} \cdots \sum_{s_N} e^{-\beta(\varepsilon_{s_1} + \ldots + \varepsilon_{s_N})} = \frac{1}{N!} Z_{\text{sp}}^N
$$

So now that we have the partition function, we can calculate the average energy $\langle E \rangle$ to be

$$
\langle E \rangle = \frac{\partial}{\partial \beta}(\ln Z_I) = \frac{1}{Z_I} \frac{\partial}{\partial \beta}(Z_1) = \frac{3N}{2\beta} = \frac{3}{2} N k_B T \implies \langle \varepsilon \rangle = \frac{<E>}{N} = \frac{3}{2} k_B T
$$

where $\langle \varepsilon \rangle$ is the average energy per particle.

2.4   Equipartition theorem

> **Equipartition Theorem**
>
> Each term in the energy that is quadratic in some coordinate contributes $\frac{k_B T}{2}$ to the average energy.

Examples:

- Consider a $1D$ mass on a spring. Then its energy is given by

$$
\varepsilon = \frac{1}{2} k x^2 + \frac{p^2}{2m} \implies \langle \varepsilon \rangle = 2 \frac{k_B T}{2} = k_B T
$$

- A diatomic molecule in free space has three degrees of freedom in its momentum of center of molecule, in addition to vibrational modes and rotational energy from its angular momentun L

$$\varepsilon = \sum_{i=x,y,z} \frac{p_i^2}{2m} + \frac{1}{2}kr^2 + \frac{p_r^2}{2m} + 2\frac{L^2}{2J} \implies \langle\varepsilon\rangle = (3 + 2 + 2)\frac{k_B T}{2} = \frac{7}{2}k_B T$$

In general, if we have that the energy of a single particle is given by m different quadratic terms:

$$\varepsilon = \sum_{i=1}^m a_i q_i^2$$
$$\implies Z_{\text{sp}} = C \int \cdots \int e^{-\beta \sum_{i=1}^m a_i q_i^2} dq_1 \ldots dq_m$$
$$= C \prod_{i=1}^m \sqrt{\frac{\pi}{\beta a_i}}$$

so if we have N particles, we get

$$Z = C'(\beta^{-m/2})^N \implies <E> = -\frac{\partial}{\partial\beta}(\ln Z) = m \cdot \frac{N}{2}k_B T \implies \langle\varepsilon\rangle = m \cdot \frac{k_B T}{2}$$

Recall from thermodynamics that we could define the **Heat capacity** as the quantity

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V$$

But when we measure the heat capacity of for example water, we see that the heat capactiy depends on the temperature. If it is solid, then we only see three degree os freedom plaing. If it is a liquid we see five and only if it is hot enough do we se all seven.

2.4.1 Maxwell Boltzmann distribution

We build a tank with one hole in it and let the molecules shoot out of it and sort them out by how far they travel. From this we can find out the probability distribution of the momentun in this particular direction.

So what is it? We see that

$$P(v_x)dv_x \propto e^{-\beta m v_x^2/2}dv_x \implies \langle v_x\rangle = 0$$

which makes sense as on average, the gasses move left and right equally likely. But so if we want to know the absolute value, we just square it and get

$$\left\langle v_x^2\right\rangle = \frac{\int_{-\infty}^{\infty} v_x^2 e^{-\beta m v_x^2/2}dv_x}{\int_{-\infty}^{\infty} e^{-\beta m v_x^2/2}dv_x} = \frac{k_B t}{m}$$

Now we can find out what the pressure of a gas is. Consider a single molecule travelling left to right between two walls with area A distance L apart. Then the pressure will be the quotient of the average force and the area. So

$$p = \frac{N}{A} \left\langle \frac{\text{momentum transfer per particle}}{\text{time between collision}} \right\rangle$$

The momentum transerred will then be $2mv_x$ and the time between the collision will be $2L/v_x$ and we get

$$
\begin{aligned}
p &= \frac{N}{A} \int_{-\infty}^{\infty} 2mv_x \cdot \frac{V}{2L} p(v_x) dv_x \\
&= \frac{m}{L} \int_{-\infty}^{\infty} v_x^2 p(v_x) dx \\
&= \frac{N}{V} m \left\langle v_x^2 \right\rangle
\end{aligned}
$$

where we found that $A \cdot L = V$ is the volume of the container. This can be rewritten to recover the ideal gas law

$$pV = Nk_BT$$

The next question is: what is the probability distribution the speed $v = |\vec{v}|$ of the particle?

$$
\begin{aligned}
P(p)dp &= \frac{4\pi V p^2}{h^3} dp \cdot \frac{e^{-\beta p^2/2m}}{Z_{\text{sp}}} \\
&= 4\pi p^2 \left(\frac{\beta}{2\pi m}\right)^{3/2} e^{-\beta p^2/2m} dp
\end{aligned}
$$

By Multiplying with N and using that $p = mv$, the number of partiles with speed between v and $v + dv$ are

$$
\begin{aligned}
n(v)dv &= 4\pi N v^2 \left(\frac{\beta m}{2\pi}\right)^{3/2} e^{-\beta mv^2/2} dv \\
&= Nf(v)dv
\end{aligned}
$$

Where $f(v)$ gives us the typical Maxwell-Boltzmann distribution of the velocity.
If we want to know the average speed we get

$$\langle v \rangle = \int_0^{\infty} vf(v)dv = \sqrt{\frac{8k_BT}{\pi m}}$$

and the peak of $f(v)$, which is the most probable speed we get

$$v_{\text{max}} = \sqrt{\frac{2k_BT}{m}}$$

2.5 Blackbody radiation

We define a **Blackbody** as an object that absorbs all light that hits it. This is again an idealized concept as most objects do reflect a small part of light.

But since they absorb everything, it must mean that if they were in thermal equilibrum with the surrounding environment, they must also *emit* some radiation.

For our model of a blackbody, we imagine a Box with some light trapped inside, having temperature T.
The Box will have one of its corner on the Origin, with dimensions L_x, L_y, L_z.
We then poke a hole in the box and add a spectrometer measuring how much light can escape.
Now what can we say about the light inside? If we imagine a metallic box, the wave equation together with the condition, that there are no charges in the box will be

$$\nabla^2 \vec{E} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} \quad \text{and} \quad \vec{\nabla}\cdot\vec{E} = 0$$

But since we als have a metal case, the boundary condition is that the parallel component of the electric field vanishes. ($E_\parallel = 0$).
In the $1D$ case we have that

$$\frac{\partial^2 \vec{E}}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}, \quad \vec{E}\cdot\hat{x} = 0, \quad E(x=0) = 0, \ E(x = L_x) = 0$$

A general solution for this is

$$\vec{E}(x,t) = \vec{E}_0 \sin(k_x x)e^{i\omega t}, \quad \text{where} \quad k_x = \frac{n\pi}{L_x}, n = 1,2,3,\dots \quad \omega = c_0 x_k, \quad \vec{E}_0 = E_0 \hat{r}$$

Each (linearly) independent solutions can be characterized by its wave-vector k_x and one of two polarisations:
Now for the $3D$ case, the general solutions will be of the form

$$\vec{E}(\vec{r},\vec{t}) = \begin{pmatrix} E_{0x}\cos(k_x x)\sin(k_y y)\sin(k_z z) \\ E_{0y}\sin(k_x x)\cos(k_y y)\sin(k_z z) \\ E_{0z}\sin(k_x x)\sin(k_y y)\cos(k_z z) \end{pmatrix} \cdot e^{i\omega t}$$

Where we have the following conditions

$$\vec{k} = \left(\frac{n_x\pi}{L_x}, \frac{n_y\pi}{L_y}, \frac{n_z\pi}{L_z}\right), \quad \text{for} \quad n_x, n_y, n_z = 0,1,2,\dots \quad \omega = c|\vec{k}|, \quad \vec{E}_0\cdot\vec{k} = 0$$

Here again each solution can be characzerized by its wave vector $\vec{k}$ and one of 8 possible polarisations.

If we recall the harmonic oscillator, its energy ε will be of the form

$$\varepsilon = \frac{1}{2}k\dot{x}^2 + \frac{p^2}{2m}$$

which can also be expressed in terms of the electormagnetic field and Volume V:

$$\varepsilon = \frac{V}{2}\left(\varepsilon_0|\vec{E}|^2 + \frac{1}{\mu_0}|\vec{B}|^2\right)$$

Using the equipartition theorem, we see that the two quadratic terms give us that the average energy per mode at temperature T is k_BT.

We can now ask what the energy of the emitted light is with frequencyy between ω and $\omega + d\omega$. We can also ask how many modes there are that have frequency between ω and $\omega + d\omega$ or between k and $k + dk$.

In the space of all possible k vectors we have that a shell with radius k and thickness dk will have the volume

$$\frac{1}{8}4\pi k^2 dk$$

And since all the modes are equally spaced, we can ascribe each mode its own volume. So the volume per mode will be

$$\frac{\pi}{L_x}\cdot\frac{\pi}{L_y}\cdot\frac{\pi}{L_z} = \frac{\pi^3}{V}$$

We also need to account that there are 2 polarisations per k vector, so the total number of modes between k and $k + dk$ will be

$$g(k)dk = \frac{1}{8}4\pi k^2 dk \cdot \frac{V}{\pi^3}\cdot 2 = \frac{V}{\pi^2}k^2 dk$$

Writing $\omega = ck$, we get the **Rayleigh-Jeans Law**

$$\rho(\omega)d\omega = k_BT \cdot \frac{g(\omega)d\omega}{V} = k_BT\frac{\omega^2}{\pi^2 c^3}d\omega$$

Which presents a problem! As ω goes to infinity, we would have that $\rho(\omega)\to\infty$ which means that we expect an infinite energy density.
This problem is also called the *Ultraviolet Catastrophy*. So how do we resolve it?

2.6   Planck distribution

What Planck proposed was that for each mode, the energy in that mode can only take on discrete values, which are non-negative integer values of a certain quality.

$$\varepsilon = n\cdot h\nu = n\hbar\omega,\quad n = 1, 2, 3, \ldots$$

where $\hbar = \frac{h}{2\pi}$ is the reduced Planck's constant.

In this case, the partition function of a mode will be

$$Z = \sum_{n=0}^{\infty} e^{-\beta n \hbar \omega} = \frac{1}{1 - e^{-\beta \hbar \omega}}$$

where the equipartition theorem gives us that the average energy will be

$$\langle \varepsilon \rangle = -\frac{\partial}{\partial \beta}(\ln Z) = -\frac{1}{Z}\frac{\partial}{\partial \beta}(Z)$$

$$= (1 - e^{-\beta \hbar \omega})\frac{\hbar \omega e^{-\beta \hbar \omega}}{(1 - e^{-\beta \hbar \omega})^2}$$

$$= \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \simeq k_B T$$

when $k_B T \gg \hbar \omega$ so using the Planck distribution we get

$$\rho(\omega)d\omega = \frac{\hbar \omega^3}{\pi^2 c^3}\frac{d\omega}{e^{\beta \hbar \omega} - 1}$$

which behaves normally in the sense that for $\omega \to \infty$ we have $\rho(\omega) \to 0$ (See Steve 3.15). Here the curve has a peak at

$$\rho_{\max}(\nu) \text{ is at } \hbar \omega \simeq 2.822 k_B T$$

Which is called the *Wien displacedment Law.*
From the Plank distribution, we can calculate the total Energy density u, given by the **Stefan-Boltzmann Law**

$$u(T) = \int_0^{\infty} \frac{\hbar \omega^3}{\pi^2 c^3}\frac{d\omega}{e^{\beta \hbar \omega} - 1} = aT^4$$

where a is the constant $a = \frac{\pi^2 k_B^4}{15 \hbar^3 c^3}$
This allows us to describe the Power radiated per unit area of the blackbody surface:

$$j(T) = u(T)\frac{c}{4} = \sigma T^4$$

where $\sigma = \frac{ac}{4} = 5.67 \times 10^8 \frac{W}{m^2 K^2}$ is called *Stefan's constant.*

Now we want to see what the energy inside an infinitesimal volume looks like:

$$dE = R^2 \sin \theta d\theta d\varphi dr \cdot u(T)$$

then de energy that will make it out of of the small hole dS is

$$d\varepsilon = dE\frac{dS \cdot \cos \theta}{4\pi R^2}$$

And the energy in the volume that moves through dS in a tmie dt, with $\frac{dr}{dt} = c$. So the total power radiated will be

$$j(T) = \frac{1}{dS}\frac{d\varepsilon}{dt} = \frac{cu(T)}{4\pi} = \int_0^{\pi/2} d\theta \int_0^{2\pi} d\varphi \cos\theta \sin\varphi = \frac{1}{4\pi}\frac{1}{2}2\pi cu(T) = \frac{1}{4}cu(T)$$

In order to better under stand matter better, let's first look at some key facts: Atoms are made up

of

	Charge	Mass
Protons	$1.6 \cdot 10^{-19}$ Coulomb	$1.66 \cdot 10^{-27}$ k g
Neutrons	0	$1.66 \cdot 10^{-27}$ k g
Electrons	$-1.6 \cdot 10^{-19}$ k g	$9.11 \cdot 10^{-31}$ k g

The different atoms can be ordered in the periodic table, where many of their properties can be seen, such as their atomic mass, (which is not quite an integer since atoms can have multiple *isotopes*) and the number of protons (which determines the name we give an atom).

For example, Carbon has 6 protons and electrons, but there are two isotopes. One with 6 neutrons and one with 7.

So how can we tell different isotopes apart? We can find out their atomic mass, by making a gas and using the ideal gas law $pV = Nk_BT$ and measuring the pressure, volume and temperatur, we can infer N and after weighing the gas, we can find out the average mass per atom.

This method is obviously pretty unrealistic, as you would need a gigantic amount of them.

When calculating these masses, we found that the mass is about an integer multiple of the mass of a Hydrogen atom.

But what's bad about using Hydrogen as the ground base is that it is very reactive, so people have opted to base it around the carbon atom and defined the **atomic mass unit** $u := \frac{1}{12}$-th of the mass of a carbon atom.

A better way of measuring the mass of an atom is to heat up the atoms and shoot the ion beam through a canal where we have an electromagnetic field and measure how much it gets redirected by the field.

The force on an ion is going to be given be the Lorentz equation

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$$

The effect of the electiv field will then be

$$m\frac{d^2y}{dt^2} = qE \implies y = \frac{qEt^2}{2m} \simeq \frac{qEl^2}{2mv^2}$$

where used the approximation that the velocity in the forward directionn will not change much, as the canal will be short.

The effect of the magnetic field will be

$$m\frac{d^2x}{dt^2} = -qvB \implies x = -\frac{1vB}{2m}t^2 \simeq -\frac{qBl^2}{2mv}$$

Next we want to elminitate the velocity from the equation by writing out the relationship between the x and y deflections and we get

$$y = \frac{m}{q}\frac{2E}{B^2l^2}x^2$$

which is a parabola! So by looking where the the point lies on the parabola, regardless of its velocity, we can find out its mass, assuming we know q, E, B, l.

Now after we can measure the mass of atoms, we want to know how we can measure the size of atoms. This turns out to not be as simple as placing a ruler beside it as we can't really create a ruler (which is made of atoms) small enough.

The idea is that we measure the size by throwing other things at it and seeing what happens.

We could, for example, throw other atoms, electros or an α-particle (nucleus of helium) at it.

How close can atoms get to each other? By using Van der Waals equation

$$(P + \frac{q}{V^2})(V - b) = Nk_BT$$

By measuring the relations between P, V, N, T we can sketch out what a and b look like, which tells us the size of an atom.

Another way is by shining X-rays onto the atom and looking at the diffraction pattern to determine the size of the atom.

Another, more modern way is called **scattering**. The basic idea is that we heat up the gas in an oven and open a small hole in the oven, so that the gas can escape through a chamber with another gas in it, called the scattering chamber. By varying the length of the second gas in the chamber and by measuring how much gas makes it throgh we can find out how much area the second gas takes away from the path to the detector which gives us the size of the atoms.

The area that will be covered is called the *scattering cross section* and is $\gamma = \pi(r_1 + r_2)^2$, where the gas beam will be deflected.

Then we ask what is the probabilty that an atom in the beam will get scatted by an atom in the chamber when it flies trough? It will be

$$P_s = \frac{\text{total cross section of all atoms in } Adx}{A}$$
$$= n\sigma dx$$

so $N \cdot P_s$ atoms will get scattered. So the change in the number of atons will be

$$dN = -nP_S = -n\sigma N dx$$

so after the distance x the number of remaining atoms will be

$$N(x) = N_0 \cdot e^{-n\sigma x}$$

What we found from these measurements was that the size of the atom was on the order of a few Anstroms, ($\simeq 10^{-10}$ m).

The problem with this measurement was that we were assuming that the atoms were ball shaped or had a hard shell around them, which was unsure at the time.

To probe further, we need to find out the internal structure of an atom.

Some tested the scattering experiment with electrons. And from the scattering equation above we expect that an electron can travel for 10^{-7} m without being disturbed.

But what we found instead was the the electrons could travel centimeters before betting scattered by a significant amount. Moreover, we found that faster electrons were less likely to scatter than slower moving electrons.

Another experiment was done be Rutherford. Instead of shooting electrons into a gas, he shot α particles at a thin gold foil. These particles will then be scattered by the foil and we will measure the scattering distribution.

If b is the height above which the electron flies above the gold atom and we denote θ to be the angle of deflection and Φ to be the angle from the gold atom to the electron, when we have

$$\vec{F} = \frac{2ze^2}{4\pi\varepsilon_0 r^2}\hat{r} = \frac{k}{r^2}\hat{r}$$

So since the force is always radial, it means that the angular momentum $\vec{L} = m\vec{r} \times \vec{r}$ will be preserved. So

$$L_{\text{init}} = mv_0 b = m\dot{\Phi}r^2 = L_{\text{trajectory}} \implies \frac{1}{r^2} = \frac{\dot{\Phi}}{v_0 b},$$

Which gives us

$$\vec{F} = \frac{k\dot{\Phi}}{v_0 b}\hat{r} \quad \text{and} \quad \vec{F}_\perp = \frac{k\dot{\Phi}}{v_0 b}\sin\Phi$$

$$\implies m\dot{v}_\perp = \frac{k\dot{\Phi}}{v_0 b}\sin\Phi$$

so we can integrate this term from the initial time to som time in the very future where the particle is infinitely far away. Since we are dealing with conservative foreces, the final kinetic energy will the qual the initial kinetic energy, so

$$v_\perp(t_2) = v_0 \sin\theta$$

and at time t_2 we have $\Phi = \pi - \theta$. So by integrating with respect to time we get that

$$\int_{t_1}^{t_2} \frac{dv_\perp}{dt}dt = \int_{t_1}^{t_2} \frac{k}{mv_0 b}\sin(\theta)\frac{d\Phi}{dt}dt$$

$$\implies \int_0^{v_0 \sin\theta} dv_\perp = \int_0^{\pi-\theta} \frac{k\sin\Phi}{mv_0 b}d\Phi$$

So we can find out how far apart the particle was, with respect to the measured angle

$$v_0 \sin\theta = \frac{k}{mv_0 b}(1 - \cos(\pi - \theta)) = \frac{k}{mv_0 b}(1 + \cos\theta)$$

$$\implies b = \frac{k}{mv_0^2}\cot\frac{\theta}{2}$$

Now we can ask how the angle changes, as the height of the path with respect to the gold atom was. I the flux of particles per Area A is $F = \frac{R}{A}$, then

$$dR_B = F \cdot 2\pi b\,db$$

$$db = -\frac{k}{2mv_0^2 \sin^2 \frac{\theta}{2}}d\theta$$

$$dR_b = dR_\theta = \pi F \left(\frac{k}{mv_0^2}\right)^2 \frac{\cos \frac{\theta}{2}}{\sin^3 \frac{\theta}{2}}d\theta$$

Where dR_Ω is the rate of particles going trough $d\Omega$. So what is the rate of particles being scattered in the detector solid angle $d\Omega$ for $d\Omega_\theta = 2\pi \sin \theta d\theta$? We then have

$$\frac{dR_\Omega}{dR_\theta} = \frac{d\Omega}{d\Omega_\theta}$$

$$dR_\Omega = \frac{dR_\theta}{d\Omega_\theta}d\Omega$$

$$= \pi F \left(\frac{k}{mv_0^2}\right)^2 \frac{\cos \frac{\theta}{2}}{\sin^3 \frac{\theta}{2}} \frac{d\Omega}{2\pi \sin \theta}$$

$$= \frac{F}{4} \left(\frac{k}{mv_0^2}\right)^2 \frac{1}{\sin^4 \frac{\theta}{2}}d\Omega$$

Now if we have a differential cross section $d\sigma$, then the rate at which particles scatter of of $d\sigma$ will be $dR = F d\sigma$. So for the rutherford scattering we will have

$$\boxed{d\sigma = \frac{1}{4}\left(\frac{k}{mv_0^2}\right)\frac{1}{\sin^4 \frac{\theta}{2}}d\Omega}$$

So by sending N α-particles in the area A, n atoms in the target area A. THen the number of particles scattered into $d\Omega$ will be $\frac{nN}{A}d\sigma$.

So the Rutherford scattering shows us that if we keep making the target area smaller, when does the measured amount follows this prediction. So this gives us a bound on how big the nucleus is. If the measurement follows the prediction, this tells us that the coulomb potential is the only thing scattering the particles. If the measurement deviates, this is telling us that there must be other forces at play, i.e. the particles hitting the nucleus.

2.7   Photons

We saw that the Blackbody gave us some energy of radiation which comes in bundles of $hf = \hbar\omega$.

In an experiment, we charged up a plate with electrons. Then we shined light on it and we saw that the charge decreases until the plate is neutral again. If we try this again but we put plastic

glass in front of the plate, we see that this time, the plate does not get discharged.

If the try the same experiment again, but the charge is positive, then nothing happens.

We call this the **photoelectric effect**, where Photons have the ability to "remove" electrons from atoms.

In another experiment, we create a circuit consisting of two metal plates, on of which gets shined on with light. We then can measure the amount of electrons displaced by measurig the current.

We start out with yellow light and measure some voltage. If we go to green light, the voltage increases a bit. Then blue light gives an even higher voltage, with ultraviolet light giving an even higher voltage.

We also see that the speed at which the voltage reaches the maxium voltage depends on the intensity of the light.
If we plot the frequencies with their voltages, we see a linear correspondence. If we were to continue this linear relationship until the frequency $\nu = 0$, then we would write

$$E_{\max} = h\nu - \Phi$$

where Φ is called the **work-function**

Einsteins explained the photoelectric effect as follows: The energy contained in light can only be absorbed by multiples of $h\nu$ and he called this quantum of light the **photon**.
From this Φ can be seen as the energy needed to free electrons.

The photoelectric effect allows us to convert light into energy, with one method being the **photomultiplier**, where the ejected electron is accelerated by a strong electric field, crashing into a metal plate which causes more electrons to be ejected and further accelerated.
We can also look at the inverse photoelectric effect which is used in x-ray scanning for example, where electrons get smaehd into an anode, which ejects xray photons.

2.8   Rayleigh-Scattering

We can approximate the binding potential of an electron in an atom of second order and get the quadratic terms in the taylor expansion. From this we just get a harmonic oszillator

$$m\frac{d^2x}{dt^2} + m\omega_0^2 x = -eE\cos(\omega t)$$

which has the solution

$$x = x_0 \cos(\omega t), \quad \text{where} \quad x_0 = \frac{eE}{m(\omega^2 - \omega_0^2)}$$

Which creates an oszillating dipole moment p given by

$$p = p_0 \cos(\omega t) \propto x_0 \cos(\omega t)$$

This in turn generates an electric field that gets emitted

$$E_{\text{emitted}} \propto \omega^2 p_0 = \frac{\omega^2 e E}{m(\omega^2 - \omega_0^2)} \cos(\omega t)$$

Considering that the aproximation only holds for small deviations, we can assume that for our purposes $\omega \ll \omega_0$, which means

$$E_{\text{emitted}} \propto \omega^2 \cos(\omega t)$$

Missing Lecture 13

We saw that we could associate a particle with its *de Broglie Wavelength*. But how exactly can the behaviour of particles be described as a wave?

Ideally we want to define somthing like a **wave-packet**, which is a localized collection of a wave that exists in some region in space. To do this we first model a particle travelling in free 1D space with position $x = vt$. As a monochromatic plane wave we could write

$$A(z,t) = A\cos(kx - \omega t) = A\text{Re}[e^{i(kx-\omega)}]$$

Where the particle would be delocalised, as it exists in a streched region in space. To resolve this difference between wave and particle behaviour, we will use superposition of multiple waves to achieve localisation. Recall that the intensity profile from the diffraction pattern of light going through some aperture with size a was

$$I \propto |\int_{-a/2}^{a/2} dx \int_{-a/2}^{a/2} e^{-ik(Xx+Yy))}|^2$$

And if $a \gg 1/k$, then this integral averages out expect when $X = Y = 0$, where we recover Ray-optics again.

So if we take the superposition of waves with different k values, then we can get something that behaves how we want a wave packet to behave.

For example if we take $k_0 - \Delta k < k_0 + \Delta k$, then it's associate Frequency willl be

$$\omega(k) = \frac{E}{\hbar}$$

where $E = \frac{p^2}{2m}$ is the kinetic energy. And since $p = \frac{h}{\lambda} = \hbar k$, and $k = \frac{2\pi}{\lambda}$ we can write

$$\omega = \frac{\hbar k^2}{2m}$$

If we assume a small range of k values (and momentum values): $\Delta k \ll k_0$ we can Taylor expand the frequency

$$\omega = \omega(k_0) + \frac{\partial \omega}{\partial k}\Big|_{k_0}(k - k_0) + \ldots = \frac{\hbar k_0^2}{2m} + \frac{\hbar k_0}{m}(k - k_0)$$

And by adding up all the waves (over the continuous variable k) we get that the resulting wave will be

$$\psi(x,t) = \int_{k_0-\Delta k}^{k_0+\delta k} A e^{i(kx-\omega(k)t)} dk$$

$$\simeq A e^{i\left(\frac{\hbar k_0^2}{m}t - \omega_0 t\right)} \int_{k_0-\Delta k}^{k_0+\Delta} e^{ik\left(x-\frac{\hbar k_0}{m}t\right)} dk$$

$$= 2A e^{i(k_0 x - \omega_0 t)} \operatorname{sinc}\left[\Delta k \left(x - \frac{\hbar k_0}{m}t\right)\right]$$

which reminds us of the diffraction pattern mentioned before. Recall that we cound generalize our square aperture using a transmission function τ, if we do the same here, we could write our boundary $k_0 - \Delta k < k < k_0 + \Delta k$ as a function in the k-space:

$$\psi(\tilde{k},t) = \begin{cases} A e^{-i\omega(k)t} & \text{for } |k-k_0| < \Delta k \\ 0 & \text{else} \end{cases}$$

Just like with the fourier transform, where we can either describe a function through the values it takes in the "real space" or describe it through its fourier transform which describes the frequencies of the function, we can describe wave packets either directly through the function $\psi(x,t)$ or through its frequencies $\psi(\tilde{k},t)$ in the k-space.

Now let's find out the relationship between the fourier transform and the sinc envelope. The sinc envelope has a peak at

$$x = x_{\max} = \frac{\hbar k_0}{m}t$$

which allows us to define the **group velocity** v_0 as

$$v_0 := \frac{dx_{\max}}{dt} = \frac{\hbar k_0}{m} = \frac{p_0}{m}$$

which corresponds to the classical velocity of a moving particle. The width of the sinc envelope Δx corresponds to one period of sin in sinc, so

$$\Delta k \Delta x = 2\pi, \implies \Delta p \Delta x = \hbar \Delta k \Delta x = 2\pi\hbar = h$$

where Δk is the width of $\psi(\tilde{k},t)$.

So what does $\psi(x,t)$ tell us about physical measurements? If we see $P(x,t)dx$ as the probability of finding the particle between x and $x + dx$, then

$$P(x,t)dx = |\psi(x,t)|^2 dx \quad \text{where} \quad |\psi(x,t)|^2 = \psi^*(x,t)\psi(x,t)$$

So since all the probabilities have to add up to 1 we must have that by Parsevals theorem that

$$\int_{-\infty}^{\infty} |\tilde{\psi}(k,t)|^2 dk = \int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = \int_{-\infty}^{\infty} P(x,t)dx = 1$$

Therefore we can find out that the constant A must satisfy $|A|^2 = \frac{1}{2\Delta}$.

And since $\psi(x,t)$ describes a single particle, it means that particles can interfere with themselves. In an experiment, where we shoot single alpha particles through a double slit, we get an interference pattern, where

$$\sin\theta = \frac{\lambda}{\Delta x} = \Delta p \frac{x}{p} \implies \Delta p \Delta x = \lambda p = h$$

Using the standard deviation in x and p we see that this is consistent with the Heisenberg uncertainty relation

$$\Delta x^2 = \left\langle x^2 \right\rangle - \left\langle x \right\rangle^2 \implies \Delta px \Delta x \geq \frac{\hbar}{2}$$

3 Quantum Mechanics

In the past lectures we saw many experiments which do not fit with classical mechanics, where we are working under the assumption that matter consists of particles and that we could measure properties of a particle without any concern of how the particle will be affected.

Quantum mechanics was created to reconcile the classical view with the experiments covered in the past sections. It is a *mathematical tool* to help us predict the behaviour of the objects we consider.

3.1 The wavefunction

In quantum mechanics, matter is described as particle waves, where the state of a quantum system is completely described by the **wavefunction**, which is a complex valued function with spacetime as its domain.

In the case where we have 1D space, we will write $\psi(x,t)$ which has the following properties.

1. The probability of finding the particle at the time t between x and $x + dx$ is $|\psi(x,t)|^2 dx$

2. The probability density will also have the *normalisation condition*

$$\int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = \int_{-\infty}^{\infty} P(x,t)dx = 1$$

For example, if we consider a free particle with momentum $p_0 = \hbar k_0$ this can be described by

$$\psi(x,t) = Ae^{i(k_0 x - \omega_0 t)}, \quad \text{for} \quad \hbar w_0 = \frac{p_0^2}{2m}$$

Notice that this particle has perfectly described momentum, but in turn its position is spread out through space, which makes it very non-physical. Also notice that the constant A is determined by the normalisation condition.

A useful tool when describing wavefunctions are **basis sets of functions**. Which is a collection of functions (φ_n) that are orthonormal:

$$\int_D \varphi_n^*(x)\varphi_m(x)dx = \delta_{mn}$$

and such that any function $\psi(x)$ can be written as a linear combination of these vectors.

$$\psi(x) = \sum \psi^n \varphi_n(x) \quad \text{where} \quad \psi^n = \int_D \varphi_n^*(x)\psi(x)dx$$

, since from the linearity of the integral we have that

$$\int_D \varphi_n^*(x)\psi(x)dx = \int_D \varphi_n^*(x) \sum_m \psi^m \varphi_m(x)dx$$
$$= \sum \psi^m \underbrace{\int_D \varphi_n^*(X)\varphi_m(x)dx}_{\delta_{mn}} = \psi^n$$

Now in order to show that the set of functions is complete, it must be that

$$\psi(x) \stackrel{!}{=} \sum_n \left[\int_D \varphi_n^*(x')\psi(x')dx' \right] \varphi_n(x)$$
$$= \int_D \left[\sum_n \varphi_n^*(x')\varphi_n(x) \right] \psi(x't)dx'$$

Recall that the dirac delta distribution has the property

$$\int \delta(x - x')g(x')dx' = g(x)$$

so in order for the equation to hold, it must be that

$$\sum_n \varphi_n^*(x')\varphi_n(x) = \delta(x - x')$$

Which gives us the **completeness condition** for the set of functions (φ_n).

An example of an orthonormal basis of functions is given in the Fourier series. For the Domain $\left(-\frac{L}{2}, \frac{L}{2} \right)$ we can write a function $\psi(x)$ as

$$\psi(x = \frac{1}{\sqrt{L}} \sum_n \psi^n e^{\frac{i2\pi n}{L}x}$$

Where the factor $\frac{1}{\sqrt{L}}$ is needed such that

$$\int_{-L/2}^{L/2} |\varphi_n(x)|^2 dx = 1$$

To find the expansion coefficients we can calculate

$$\int_{-L/2}^{L/2} \psi(x)\varphi_m^*(x)dx = \frac{1}{L} \sum_n \psi^n \int_{-L/2}^{L/2} e^{\frac{2\pi(n-m)}{L}x}dx$$
$$\text{for} \quad \frac{1}{L} \int_{-L/2}^{L/2} e^{\frac{i2\pi(n-m)}{L}x}dx = \delta_{mn}$$

If the dimension of our vector space is large enough, we can label the basis through a continuous variable α. Here the orthonormality of the basis vector is described in terms of the dirac delta:

$$\int \varphi^*(\alpha, x)\varphi(\alpha', x)dx = \delta(\alpha - \alpha')$$

So instead of having a sum over the coefficients, we integrate

$$\varphi(x) = \int \psi \tilde{(}\alpha)\varphi(\alpha, x)d\alpha$$

The the "coefficients" can be determined by

$$\psi(\alpha) = \int \varphi^*(\alpha, x)\psi(x)dx$$

3.2 Operators

Operators represent observables, which are procedures that are "applied" to functions. For example on the free particle

$$\psi_k(x, t) = Ae^{i(kx-\omega t)}$$

we can use the identity operator $\hat{I}$ which multiplies the function by 1.
Another operator is the one that takes the derivative with respect to x, so

$$\frac{\partial}{\partial x}\psi(x, t) = ikAe^{i(kx-\omega t)} = \frac{ip}{\hbar}\psi(x, t)$$

notice that the operator returned a scalar multiple of the function. In this case we call $\frac{ip}{\hbar}$ an **eigenvalue** and the free particle $\psi(x, t)$ an **eigenfunction**. Also notice that we were able to obtain the momentum from this operator. So we can modify it slightly to get the **momentum operator** $\hat{P} = -i\hbar\frac{\partial}{\partial x}$.
When we want to measure the position of a particle, we can use the **position operator** $\hat{x} = x$. Now what are the Eigenfunctions for this operator? We can find this out by solving

$$\hat{x}\psi_\alpha(x) = x\psi_\alpha(x) \stackrel{!}{=} \alpha\psi_\alpha(x)$$

Which can only be satisfied by the dirac delta functions.

$$\psi_\alpha(x) = A\psi(x - \alpha)$$

It turns out that every physical, measurement is represented by **hermitian** operators, whose eigenvalues are the possible measureable results. Recall from Linear Algebra that the hermitian operators have real eigenvalues.
That is if you measure an observable, and the system has a wavefunction that is an eigenfunction of the corresponding operator, then the result will always be the eigenvalue.

3.3   Change of basis

We know that the eigenfunctions of the momentum operator $\hat{p}$

$$\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{ip}{\hbar}x}$$

aswell as the Eigenfunctions of the position operator $\hat{x}$

$$\varphi_\alpha(x) = \delta(x - \alpha)$$

Form basis of the function space. This means that we can write any function using a linear combinaion of these basis functions. If we want to find out the expansion coefficients in terms of the eigenfunctions of $\hat{x}$, then we get

$$\tilde{\psi}(\alpha, t) = \int \delta(x - a)\psi(x,t)dx = \psi(\alpha, t)$$

and in terms of the eigenfucntions of $\hat{p}$ we obtain the coefficients by through

$$\tilde{\psi}(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int e^{\frac{-ip}{\hbar}x}\psi(x,t)dx$$

Notice that both representation gives us the exact same information about the particle, since the coefficients $\tilde{\psi}(p,t)$ and $\tilde{\psi}(\alpha,t)$ give us the original function $\psi(x,t)$ or its fourier transform $\tilde{\psi}(p,t)$ back.

If we want to know the probability of measuring the system between x and $x + dx$, then we want to use the expansion with respect to $\hat{x}$, since

$$P(x)dx = |\psi(x,t)|^2dx$$

and if we ant to know the probabilty of measuring the momenum between p and $p + dp$, then we use the expansion with respect to $\hat{p}$

$$P(p)dp = |\tilde{\psi}(p,t)|^2dp$$

In general for any other basis, we obtain the following theorem.

Spectral Theorem

The Eigenfunctions of an operator $\hat{O}$ form a complete, orthonormal basis and the probability of measuring an observable to be between O and $O + dO$ is

$$P(O)dO = |\psi_O(O,t)|^2dO$$

Lecture 16 Missing

3.4   Schrödinger equation

Since the wavefunction is also dependent on time, we want to know how the wave function evolves over time. The **Schrödinger Equation** does just that $\Psi(x,t)$. In the 1D case, the Schrödinger equation can be written in its **position representation**[1]:

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi$$

where $V(x)$ is the potential energy function. Here we only consider conservative forces $F = -\frac{\partial V(x)}{\partial x}$. Notice that the right hand side of the Schrödinger Equation is also linear Operator acting on Ψ. We call this the the **Hamiltonian operator** $\hat{H}$, which is given by

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \hat{V}(x) = \frac{\hat{p}^2}{2m} + \hat{V}$$

which allows us to rewrite the Schrödinger equation as

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$$

Note that the Hamiltonian measures the sum of kinetic and potential Energy. It is also a hermitian operator, whose real Eigenvalues represent observables that are the definite Energies of the corresponding Eigenstates.

We write $\Psi_n = \psi_n(x)\tau_n(t)$ for the Eigenstates of $\hat{H}$. As we've seen, these Eigenstates form a basis and they satisfy the Eigenvektor equation, which we call the **time-independent Schrödinger Equation** (TISE):

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

We can solve the Eigenvektors of the Hamiltonian, which is just a first order linear PDE, to get the solution

$$E_n\Psi_n = \hat{H}\Psi_n = i\hbar\frac{\partial}{\partial t}\Psi_n$$
$$\implies \Psi_n(x,t) = e^{-iE_n t/\hbar}\psi_n(x)$$

in terms of the initial state $\psi_n(x) := \Psi_n(x, t = 0)$.

Notice that the exponential part has asbsolute value 1, which means that the probability densitiy does not change in time:

$$|\Psi_n(x,t)|^2 = \psi_n^*\psi_n e^{iE_n\frac{t}{\hbar}}e^{-iE_n\frac{t}{\hbar}} = |\psi_x(n)|^2 = |\Psi_n(x,t=0)|$$

Since these energy Eigenstates form a basis, we can solve the Schrödinger Equation using this recipe:

---

[1]as opposed to its **momentum representation** $\Psi(k,t)$

Solving the Schrödinger Equation

1. You are given a potential $V(x)$ and an initial state $|\Psi_0\rangle$

2. Solve the time-independent Schrödinger Equationn

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

for the Eigenstates $|\psi_n\rangle$, and their Eigenenergies E_n.

3. Write the initial state of of the wave function $|\Psi_0\rangle$ in terms of the basis of initial Eigenstates $|\psi_n\rangle$. When the Eigenvalues are discrete, we need to find the coefficients C_n such that

$$|\Psi_0\rangle = \sum_n C_n |\psi_n\rangle$$

If the Eigenvalues are continuous, take the integral:

$$\text{missing}\#\#\#$$

4. The solution to the Schrödinger Equation can is then the linear combination of the time-developed Eigenstates ψ_n:

$$|\Psi(t)\rangle = \sum_n C_n |\Psi_n\rangle = \sum_n C_n e^{-iE_n t/\hbar} |\psi_n\rangle$$

Note that this works because the wave equation is linear, so a linear combination of solutions gives us another solution.
Also note that dirac notation is used, which means that we can chose to any basis we want to find the solutions.
In the position representation, this would give us

$$\langle x|\Psi(t)\rangle = \sum_n C_n e^{-iE_n\frac{t}{\hbar}} \langle x|\psi_n\rangle$$

$$\implies \Psi(x,t) = \sum_n C_n e^{-iE_n\frac{t}{\hbar}} \psi_n(x)$$

, where we used that the position operator is linear.

Now let's solve some Schrödinger Equations using this method:

3.4.1   Free particle

In our first example, we consider the simplest scenario: that of a free particle. Here, the potential is set to zero so the Hamiltonian only consists of the kinetic energy part:

$$\hat{H} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$$

In this case, the energy Eigenstates are just the momentum Eigenstates

$$\hat{H} \ket{p} = \frac{p^2}{2m} \ket{p}$$

Note that in this case, both $\ket{p}$ and $\ket{-p}$ have the same energy, so if we were to make a precise measurement we would observe the energy $\frac{p^2}{2m}$, the state collapses to a superposition of both such states:

$$\ket{\Psi} = A \ket{p} + B \ket{-p}$$

Or, since $p = \hbar k$, in the position representation:

$$\Psi(x,t) = A e^{ik\left(x - \frac{\hbar k}{2m} t\right)} + B e^{-ik\left(x + \frac{\hbar k}{2m} t\right)}$$

Because we have that multiple eigenstates can have the same eigenvalue, we say that the operator $\hat{H}$ is **degenerate** here. It has the consequence where a measurement of an observable does not collapse the wave function into a single eigenstate.

Usually we would associate the **velocity** of a wave with $v = \frac{\omega}{k}$, which in this case would be $v = \frac{\hbar k}{2m}$. However, in the classical sense of velocity we should be $v' = \frac{p}{m} = \frac{\hbar k}{m} = 2v$.
We can resolve this issue by saying that v is the **phase velocity**, the velocity at which the phase front moves. But since the wave function is not localized, the wave is not going anywhere so the concept of movement doesn't really apply here as our particle is very un-physical.
In more physical settings, we write $v_{\text{group}} = \frac{\partial \omega}{\partial k}$ for the **group velocity** of a particle, which matches with the classical notion of velocity.

3.4.2   The infinite square well

Here we consider a particle that is "trapped" inside a well of infinite potential $V(x)$, which is given by

$$V(x) = \begin{cases} 0 & 0 \leq x \leq a \\ \infty & \text{elsewhere} \end{cases}$$

Since the potential outside of the well is infinite, the wave function must be zero or else its energy would be infinite. Inside the well, we have the same scenario as in the free particle, so we can write

$$\begin{aligned} \Psi(x) &= C_+ e^{ikx} + C_- e^{-ikx} \\ &= A \cos(kx) + B \sin(kx) \end{aligned}$$

In order to fix the constants A, B, we need more conditions for the wave function.
We do this by imposing **boundary conditions** to our problem, which are motivated by our physical observations. In particular, we want

1. Continuity of the wave function $\Psi(x)$ in x.

2. Continuity of the derivative $\frac{\partial \Psi(x)}{\partial x}$, except where $V(X) \to \infty$. We can motivate this by integrating the Schrödinger Equation over a region from $-\varepsilon$ to ε and taking the limit

$$\lim_{\varepsilon \to 0} \frac{\partial \psi}{\partial x}\Big|_{+\varepsilon} - \frac{\partial \psi}{\partial x}\Big|_{-\varepsilon} = \lim_{\varepsilon \to 0} -\frac{\hbar^2}{2m} \int_{-\varepsilon}^{\varepsilon} \frac{\partial^2 \psi}{\partial x^2} dx + \lim_{\varepsilon \to 0} \int_{-\varepsilon}^{\varepsilon} V(x)\psi(x)dx$$
$$= \lim_{\varepsilon \to 0} E \int_{-\varepsilon}^{\varepsilon} \psi(x)dx$$

which is zero if E is finite and ψ is continuous.

By applying the first continuity condition, we must have that at the first wall, $(x = 0)$ we have that ψ must satisfy

$$\psi(0) = A\sin(0) + B\cos(0) = B = 0$$

and for the second Wall $(x = a)$:

$$\psi(a) = A\sin(ka) = 0 \implies ka = n\pi, \text{ for } n \in \mathbb{Z}$$

For $n = 0$ this gives us the trivial solution $\psi = 0$, also since the sine is an odd function, it suffices to look at positive n, which we will label the discrete eigenfunctions

$$\psi_n(x) = A\sin(k_n x) \text{ for } k_n = \frac{n\pi}{a}$$

Now in order to find the the factor A, we must test that the integral over the probability distribution must equal 1,so

$$\int_0^a |\psi_n(x)|^2 dx = \int_0^a |A|)^2 \sin^2(k_n x) dx = |A|^2 \frac{a}{2} = 1 \implies A = \sqrt{\frac{2}{a}}$$

so we can write the eigenstate as

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin(k_n x) \text{ for } k_n = \frac{n\pi}{a}$$

Next we need to find out the energy E_n of the eigenstate $|\psi_n\rangle$. It is given by

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

Before we move on to write the initial state Ψ_0 in terms of these eigenstates, we note the following:

- The energies are **quantized**, as n can only take integer values.

- The eigenstates are just the basis function for the fourier sine series expansion.

- If we graph these eigenfunctions, we see that they have n peaks/valleys. As n increases, the second derivative increases and therefore the kinetic energy increases.

- The eigenfunctions alternate between even and odd functions relative to the center of the well.

  If we use shifted coordinates $y = x - \frac{a}{2}$, such that the well is centered at the origin we have

  $$\psi_n(y) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}(y + \frac{a}{2})\right)$$

  $$= \sqrt{\frac{2}{a}} \begin{cases} \cos(\frac{n\pi}{a}y) & \text{for } n \text{ odd} \\ -\sin(\frac{n\pi}{a}y) & \text{for } n \text{ even} \end{cases}$$

The last point brings us the concept of **symmetry**, which can be used to simplify our problem: Since the potential V is symmetrical, we can introduce the **parity operator** $\hat{\Pi}$ given by

$$\hat{\Pi}\psi(y) = \psi(-y)$$

which commutes with the Hamiltonian $\hat{H}$, since

$$(\hat{\Pi} \circ \hat{H})\psi(y) = \hat{H}\psi(-y) = (\hat{H} \circ \hat{\Pi}\psi(y) \implies [\hat{H}, \hat{\Pi}] = 0$$

Which means that both $\hat{H}$ and $\hat{\Pi}$ share the same eigenfunctions φ_π. Now if we were to act the parity operator twice, we obtain the identity operator $\hat{I}$. Therefore, the Eigenvalues have to be ± 1, where the positive Eigenvalue $+1$ corresponds to an even function and -1 corresponds to odd functions.

This means that before solving the equation, we could have found out that the solution is a superposition of even and odd functions.
More generally, the symmetries can give us great insight into the solutions space before attempting to solve the Schrödinger equation.

3.4.3   The finite square well

In our third example, we have that our well only has a finite potential, which is given by

$$V(x) = \begin{cases} 0 & |x| \leq \frac{a}{2} \\ V_0 & \text{elsewhere} \end{cases}$$

We divide the space into three regions. We call the region on the left region I, the one in the well region II and the one on the right region III.
For region I, the Schrödinger equation is

$$\hat{H}\Psi_E = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi_E + V_0\Psi_E = E\Psi_E$$

which has solutions of the form

$$\Psi_E(x) = A_\mathrm{I} e^{-\sqrt{k_0^2 - k^2}x} + B_I e^{\sqrt{k_0^2 - k^2}x}$$

where we introduced the constants k_0, k through

$$k_0 = \sqrt{\frac{2mv_0}{\hbar^2}} \text{ and } k = \sqrt{\frac{2mE}{\hbar^2}}$$

For region II, the Schrödinger equation is that of a free particle:

$$\hat{H}\Psi_E = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi_E = E\Psi_E$$

We could solve this using the general solution like in region I, but we can also argue that the potential is symmetric, so the solutions must consist of even and odd functions:

$$\Psi_E = A_{\text{II}}\cos(kx) + B_{\text{II}}\sin(kx)$$

For Region III, we use the same general solution as in region I:

$$\Psi_E(x) = A_{\text{III}}e^{-\sqrt{k_0^2-k^2}x} + B_{\text{III}}e^{\sqrt{k_0^2-k^2}x}$$

Now we will seperately consider the two cases, where the Energy of the particle can either be bigger or smaller than the potential of the well:

$E < V_0$ In this case, we know that $k < k_0$, so $\sqrt{k_0^2 - k^2}$ is real. For the even solutions we have

$$B_{\text{II}} = 0, \quad A_{\text{I}} = B_{\text{III}}, \quad B_{\text{I}} = A_{\text{III}}$$

further, we have to normalize the function. Since $\sqrt{k_0^2 - k^2}$ is real, the exponential terms would diverge, so $B_{\text{III}} = A_{\text{I}} = 0$.

By continuity of Ψ_E at the endpoints we must have that

$$B_I e^{\sqrt{k_0^2-k^2}(-a/2)} = A_{\text{II}}\cos(-ka/2)$$

and by continuity of the derivative, at $x = -\frac{a}{2}$ we have

$$B_I\sqrt{k_0^2 - k^2}e^{\sqrt{k_0^2-k^2}(-a/2)} = -A_{\text{II}}k\sin(-ka/2)$$

and by dividing the equations we get

$$\frac{k_0^2 - k^2}{k} = \tan\left(\frac{ka}{2}\right)$$

For the odd solutions we can again argue from symmetry that

$$A_{\text{II}} = 0, \quad A_{\text{I}} = -B_{\text{III}}$$

and if we go through the same equations as above we get the solutions

$$\frac{\sqrt{k_0^2 - k^2}}{k} = -\cot\left(\frac{ka}{2}\right)$$

which concludes all cases for $E < V_0$.

$E > V_0$ In the case where the energy of the well is bigger than the Potential off the well, the particle can reach unbound states, that extend to infinity. Here, $\sqrt{k_0^2 - k^2}$ is imaginary so instead of exponential decay, the wavefunction will be periodic in the outside regions. The boundary contditions for the even solutions at $x = -\frac{a}{2}$ are

$$A_\mathrm{I} e^{i\frac{a}{2}\sqrt{k^2-k_0^2}} + B_\mathrm{I} e^{-i\frac{a}{2}\sqrt{k^2-k_0^2}} = A_\mathrm{II} \cos(ka/2)$$

$$-iA_\mathrm{I}\sqrt{k^2-k_0^2} e^{i\frac{a}{2}\sqrt{k^2-k_0^2}} + B_\mathrm{I} e^{-i\frac{a}{2}\sqrt{k^2-k_0^2}} = A_\mathrm{II} k \sin\left(\frac{ka}{2}\right)$$

The solutions for these wavefunctions are not normalizable.

In general, where we have any potential, we will have that if the energy is smaller than the highest potential, the wave function will be in a **bound state**. And if the Energy exceeds that, the particle will have continuous Eigenvalues and be delocalized.

3.4.4   The step potential

In this example, we wil look at a heaviside stepfunction

$$V(x) = \begin{cases} V_0 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

We again seperate space into two regions: Region I on the left and Region II on the right. And we consider a particle coming form th left.
The general solution in Region I will be of the form

$$\Psi_E(x) = A_I e^{ikx} + B_I e^{-ikx}$$

which corresponds to two waves that are travelling from left to right ($A_I e^{ikx}$) and one from travelling from right to left ($B_I e^{-ikx}$).
For Region II, the general solution is

$$\Psi_E(x) = A_{II} e^{\sqrt{k_0^2-k^2}x} + B_{II} e^{-\sqrt{k_0^2-k^2}x}$$

, for $k = \sqrt{\frac{2mE}{\hbar^2}}, k_0 = \sqrt{\frac{2mV_0}{\hbar^2}}$.

$E > V_0$ In the case where the Energy is higher than the potential V_0, the exponent is complex, so region II can be described as

$$\Psi_E(x) = A_{II} e^{iqx} + B_{II} e^{-iqx}$$

where $q = \sqrt{k^2 - k_0^2}$ is real. Since the particale is coming from the left, we can set B_II to be zero. The same cannot be said for B_I as it might reflect off of the potential.

From the continuity of ψ at $x = 0$ and its derivative, we obtain

$$A_\mathrm{I} + B_\mathrm{I} = A_\mathrm{II}$$

$$ikA_\mathrm{I} - ikB_\mathrm{I} iqA_\mathrm{II}$$

which we can combine to get

$$A_{\text{II}} = \frac{2k}{k+q}A_{\text{I}} \quad \text{and} \quad B_{\text{I}} = \frac{k-q}{k+q}A_{\text{I}}$$

Then we define reflection coefficients r and t as the ratios

$$|r|^2 := |\frac{B_{\text{I}}}{A_{\text{I}}}|^2 = |\frac{k-q}{k+q}|^2$$

$$|t|^2 := |\frac{A_{\text{II}}}{A_{\text{I}}}|^2 = |\frac{2k}{k+q}|^2$$

Notice that $|r|^2 + |t|^2 = 1$. And even if the energy of the particle is higher than the step height, there is always going to be some reflection.

$E < V_0$ Since the exponent $K = \sqrt{k_0^2 - k^2}$ is real, A_{II} must be zero, or else our function is not normalizable. The boundary conditions give

$$A_{\text{I}} + B_{\text{I}} = B_{\text{II}}$$

$$ikA_{\text{i}} - ikB_{\text{I}} = -KB_{\text{II}}$$

which give us the relations

$$B_{\text{I}} = -\frac{K+ik}{K-ik}A_{\text{I}} \quad \text{and} \quad B_{\text{II}} = -\frac{2ik}{K-ik}A$$

notice that despite our particle having less energy than the potential, there is a non-zero probability of finding the particle at $x > 0$.

3.5   Harmonic oscillator

One of the most important examples is when the potential is that of a harmonic oscillator.

$$V(x) = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2 x^2$$

for some *resonance frequency* $\omega = \sqrt{\frac{k}{m}}$
Notice that in this case, the Hamiltonian is quadratic in both position and momentum

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$$

What makes the harmonic oscillator nice is that in many other problems, the energy of the system can be described as a sum of two quadratic terms. For example in an electromagnetic field, we have

$$\varepsilon = \frac{1}{2}\varepsilon_0|\vec{E}|^2 + \frac{1}{2\mu_0}|\vec{B}|^2$$

Now we will introduce the dimensionless variables

$$\hat{X} := \sqrt{\frac{m\omega}{\hbar}}\hat{x} \quad \text{and} \quad \hat{P} := \sqrt{\frac{1}{\hbar m\omega}}\hat{p}$$

so we can re-write the Hamiltonian as

$$\hat{H} = \frac{\hbar\omega}{2}\left(\hat{X}^2 + \hat{P}^2\right)$$

From this representation, it becomes clear that the Hamiltonian is just a sum of two quadratic terms. Also, the factor $\hbar\omega$ characterizes the energy scale, which we also have seen to be the energy of a photon with frequency ω.

We solve this by factoring the Hamiltonian. Since $\hat{X}$ and $\hat{P}$ we have to remember the extra terms colored in orange.

$$\begin{aligned}
\hat{H} &= \frac{\hbar\omega}{2}(\hat{X}^2 + \hat{P}^2) \\
&= \frac{\hbar\omega}{2}\left[\left(\hat{X} - i\hat{P}\right)\left(\hat{X} + i\hat{P}\right) -i(\hat{X}\hat{P} - \hat{P}\hat{X})\right] \\
&= \frac{\hbar\omega}{2}\left[\hat{X}^2 + i\hat{X}\hat{P} - i\hat{P}\hat{X} + \hat{P}^2 -i(\hat{X}\hat{P} - \hat{P}\hat{X})\right]
\end{aligned}$$

If we think about the extra terms, we can show that

$$-i(\hat{X}\hat{P} - \hat{P}\hat{X}) = -i[\hat{X}, \hat{P}] = -\frac{i}{\hbar}\underbrace{[\hat{x}, \hat{p}]}_{i\hbar} = 1$$

Using this we can define two new operators. $\hat{a}$ and $\hat{a}^\dagger$, where $\cdot^\dagger$ defines the hermition conjugate.

$$\hat{a} := \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}) \quad \text{and} \quad \hat{a}^\dagger := \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P})$$

From $[\hat{X}, \hat{P}] = i$ we also find

$$[\hat{a}, \hat{a}^\dagger] = \frac{1}{2}\left([\hat{X}, -i\hat{P}] + [i\hat{P}, \hat{X}]\right) = 1$$

Which allows us to write the hamiltonian in the following way

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)$$

Suppose we had an Eigenstate $|n\rangle$ of the operator $\hat{a}^\dagger\hat{a}$, then $|n\rangle$ is also an eigenstate of the hamiltonian, since

$$\hat{a}^\dagger\hat{a}\,|n\rangle = n\,|n\rangle \implies \hat{H}\,|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle$$

Using $[\hat{a}, \hat{a}^\dagger] = 1$ and applying the operator $\hat{a}$ to both sides we get

$$\hat{a}\hat{a}^\dagger\hat{a}\ket{n} = (\hat{a}^\dagger\hat{a} + 1)\hat{a}\ket{n} = n\hat{a}\ket{n} \implies \hat{a}^\dagger\hat{a}(\hat{a}\ket{n}) = (n-1)\hat{a}\ket{n}$$

There are two was to satisfy this equation. Either trivially

$$\hat{a}\ket{n} = 0 \implies n = 0$$

or $\hat{a}\ket{n}$ is up to a constant a normalized eigenstate of $\hat{a}^\dagger\hat{a}$ with eigenvalue $n - 1$. We call this eigenstate $\ket{n-1}$ so we can write

$$\ket{n-1} = C_{n-}\hat{a}\ket{n}$$

And similarly, we can show that $\hat{a}^\dagger\ket{n}$ is another eigenstate of the operator $\hat{a}\hat{a}^\dagger$ with eigenvalue $n + 1$. So we can write

$$\ket{n+1} = C_{n+}\hat{a}^\dagger\ket{n}$$

such that $\braket{n-1|n-1} = \braket{n+1|n+1} = 1$.
This is really nice. If we were able to find an eigenstate of $\hat{a}^\dagger\hat{a}$, we are able to generate new eigenstates by applying the $\hat{a}$ and $\hat{a}^\dagger$ operators. And each time we do this, we get an eigenstate with energy higher/lower by $\hbar\omega$.
But what happens if we keep appplying $\hat{a}$ to get lower and lower energies? The only way this can be stopped if we get the case where $n = 0$. And we can no longer get a new eigensate since $\hat{a}\ket{0} = \ket{0}$. This means that the allowed energy eigenvalues of the Hamiltonian can be quantized. Since the possible n we found were the natural numbers, the energies of the Eigenstates $\ket{n}$ are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad \text{for} \quad n = 0, 1, 2, \ldots$$

We see that the energies are a set of discrete values, evenly spaced by $\hbar\omega$.
Now that we know the energies, we want to find out the spatial distribution $\psi_n(x) = \braket{x|0}$ of the eigenstates $\ket{n}$.
Starting with $n = 0$, we call $\psi_0(x) = \braket{x|0}$ the **ground state**, since it has the lowest energy. Recall that $\hat{a}\ket{0} = 0$, which we can write in the position representation as

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}) = \frac{1}{\sqrt{2\hbar m\omega}}\left(\hbar\frac{\partial}{\partial x} + m\omega x\right)$$

which gives the following condition for $\psi_0(x)$

$$\frac{\partial\psi_0(x)}{\partial x} = -\frac{m\omega}{\hbar}x\psi_0(x)$$
$$\int \frac{1}{\psi_0(x)}d\psi_0(x) = -\frac{m\omega}{\hbar}\int x dx$$
$$\implies \psi_0(x) = Ae^{-\frac{m\omega}{2\hbar}x^2}$$

for some normalisation constant A, which can be found using the normalisation condition

$$|A|^2 \int_{-\infty}^{\infty} e^{-\frac{m\omega}{\hbar}x^2} dx = |A|^2 \sqrt{\frac{\pi\hbar}{m\omega}} \overset{!}{=} 1$$

so the ground state is given by

$$\psi_0(x) = \left(\frac{\pi\hbar}{m\omega}\right)^{\frac{1}{4}} e^{-\frac{m\omega}{2\hbar}x^2}$$

And we can find the other wavefunctions with higher energy states by repeatedly applying the $\hat{a}^\dagger$ operator. To do this, we first need to find the normalisazion constant C_{n+} such that $\langle n+1|n+1\rangle = 1$.

We use that an operator $\hat{Q}$ and its hermitian conjugate $\hat{Q}^\dagger$ are adjoint pairs that satisfy $\langle f|\hat{Q}g\rangle = \langle \hat{Q}^\dagger f|g\rangle$:

$$\begin{aligned}
1 = \langle n+1|n+1\rangle &= \langle C_{n+}\hat{a}^\dagger n|C_{n+}\hat{a}^\dagger|n\rangle \\
&= \langle n|\overline{C_{n+}}C_{n+}\hat{a}\hat{a}^\dagger|n\rangle \\
&= |C_{n+}|^2 \langle n|\hat{a}\hat{a}^\dagger|n\rangle \\
&= |C_{n+}|^2 \langle n|(\hat{a}^\dagger\hat{a}+1)|n\rangle \\
&= |C_{n+}|^2(n+1)\langle n|n\rangle = |C_{n+}|^2(n+1)
\end{aligned}$$

Which gives us the following:

$$C_{n+} = \frac{1}{\sqrt{(n+1)}}, \quad |n+1\rangle = \frac{\hat{a}^\dagger}{\sqrt{n+1}}|n\rangle$$

or similarly for $\hat{a}$:

$$C_{n-} = \frac{1}{\sqrt{n}}, \quad |n-1\rangle = \frac{\hat{a}}{\sqrt{n}}|n\rangle$$

So if we have $|0\rangle$, we can find the other eigenstates with

$$|n\rangle = \frac{\left(\hat{a}^\dagger\right)}{\sqrt{n!}}|0\rangle$$

Which in our case gives us the result

$$\psi_n(X) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} H_n(x)e^{-\frac{1}{2}X^2}$$

where $X = \sqrt{\frac{m\omega}{\hbar}}x$ and H_n is the n-th **Hermite polynomial**.

3.6   Expectation values of the harmonic oscillator

We saw that we can write the Hamiltonian in terms of the "ladder operators" $\hat{a}$ and $\hat{a}^\dagger$. From this, we can also write the position and momentum operators $\hat{x}$ and $\hat{p}$ as follows

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^\dagger + \hat{a}) \quad \text{and} \quad \hat{p} = i\sqrt{\frac{\hbar m\omega}{2}}(\hat{a}^\dagger - \hat{a})$$

whose expectation values are

$$\langle n|\hat{x}|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left(\langle n|\hat{a}^\dagger|n\rangle + \langle n|\hat{a}|n\rangle\right) \quad = \sqrt{\frac{\hbar}{2m\omega}}\left(\sqrt{n+1}\langle n|n+1\rangle + \sqrt{n}\langle n|n-1\rangle\right) = 0$$

where we used the fact that the eigenstates of hermitian matrices are orthogonal. Similarly, we can show that

$$\langle n|\hat{p}|n\rangle = 0$$

Now that we know the expectation value, what about the uncertainty in x and p? We have

$$\hat{x}^2 = \frac{\hbar}{2m\omega}\left(\hat{a}^\dagger + \hat{a}^d\hat{a} + \underbrace{\hat{a}\hat{a}^\dagger}_{2\hat{a}^\dagger\hat{a}+1} + \hat{a}\hat{a}\right)$$

$$\implies \langle h|\hat{x}^2|n\rangle = \frac{\hbar}{2m\omega}(2n+1)$$

and similarly we can show that

$$\langle n|\hat{p}^2|n\rangle = \frac{\hbar m\omega}{2}(2n+1)$$

From this we can characterize the standard deviation, or spread of the wavefunction in both position and momentum space:

$$\Delta x = \sqrt{\langle \hat{x}^2\rangle - \langle \hat{x}\rangle^2} = \sqrt{\frac{\hbar}{2m\omega}}\sqrt{2n+1}$$

$$\Delta p = \sqrt{\langle \hat{p}^2\rangle - \langle \hat{p}\rangle^2} = \sqrt{\frac{\hbar m\omega}{2}}\sqrt{2n+1}$$

Notice that the ground state $|0\rangle$ has non-zero energy and non-zero uncertainty in position and momentum. We call them **zero-point energy** and **zero-point position/momentum**.
Interpreting this in the context of photons, this says that vacuum always has fluctuations on the field. We also see that the Heisenberg uncertainty is an equality for the ground state and an inequality for the excited states, since

$$\Delta x \Delta p = \frac{\hbar}{2}(2n+1)$$

3.7   Coherent states

In classical mechanics, the classical harmonic oscillator oscilllates harmonically, as expected. But how can we reconcile our findings with the classical expectation?

If we have quantum states that exhibits a classical behaviour, we call them **coherent states**, which are superpositions of energy eigenstates.

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

, where $\alpha = |\alpha|e^{i\varphi}$ is a complex number containing the amplitude and phase of the different oscillations. We can show that the expectation value for the position oscillates over time:

$$\langle\alpha|\hat{x}|\alpha\rangle = \sqrt{\frac{\hbar}{2m\omega}} \left(\alpha e^{i\omega t} + \alpha^* e^{-i\omega t}\right) \qquad\qquad = \sqrt{\frac{2\hbar}{m\omega}}|\alpha|\cos(\omega t + \varphi)$$

which also minimizes the Heisenberg uncertainty relation ($\Delta x \Delta p = \frac{\hbar}{2}$).

Another example is the light that a laser outputs, which can be well described by a coherent state.

3.8   3D Schrödinger equation

The generalisation of the Schrödinger Equation in 3D is still

$$i\hbar\frac{\partial\Delta}{\partial t} = \hat{H}\Psi$$

, but the Hamiltonian has changed a bit. It is now given by

$$\boxed{\hat{H} = \frac{1}{2m}|\hat{\boldsymbol{p}}|^2 + V(\hat{\boldsymbol{r}})}$$

where $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{p}}$ are vectors made up of operators for each spatial component

$$\hat{\boldsymbol{r}} = (\hat{x}, \hat{y}, \hat{z}) \quad \text{and} \quad \hat{\boldsymbol{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$$

In the position representation, we would write

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x}, \quad \hat{p}_y = -i\hbar\frac{\partial}{\partial y}, \quad \hat{p}_z = -i\hbar\frac{\partial}{\partial z},$$

$$\implies \hat{\boldsymbol{p}} = -i\hbar\nabla$$

So if we use the laplacian

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

we can write the schrödinger Equation as follows

$$\boxed{i\hbar\frac{\partial\Delta}{\partial t} = -\frac{\hbar^2}{2m}\Delta\Psi + V\Psi}$$

These new operators commute/ don't commute in the following way

1. For different cartesian coordinates, they commute:

$$[\hat{x}, \hat{p}_y] = [\hat{x}, \hat{p}_z] = [\hat{y}, \hat{p}_z] = \ldots = 0$$

2. Since we assume smoothness of our wavefunction, the position operators also commute by Schwarz-Clairaut's theorem.

$$[\hat{p}_x, \hat{p}_y] = [\hat{p}_y, \hat{p}_z] = [\hat{p}_z, \hat{p}_x] = 0$$

3. The position and momentum in the same catesian coordinate do not commute, just like in the 1D case

$$[\hat{x}, \hat{p}_x] = [\hat{y}, \hat{p}_y] = [\hat{z}, \hat{p}_z] = i\hbar$$

THe normalisation condition for the 3D case is now

$$\int |\Psi(\boldsymbol{r}, t)|^2 dx dy dz = 1$$

, where $\int |\Psi(\boldsymbol{r}, t)|^2 dx dy dz$ is the probability of finding the particle in the volume $dx dy dz$.
Often, the potential will be have radial symmetry, which depend solely on the distance form the origin. We call these **central potentials**.
For those it makes sense to use spherical coordinates, where the Laplacian has the form

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}$$

The symmetry means that we can seperate solutions for the problem into products of functions which depend only on the radial cordinate and the angular ones, so we can write

$$\Psi(r, \theta, \varphi) = R(r) Y(\theta, \varphi)$$

so the TISE looks like this

$$-\frac{\hbar^2}{2m} \left[\frac{Y}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{R}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{R}{r^2 \sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} \right] + VRY = ERY$$

Dividing by RY and multiplying by $-\frac{2mr^2}{\hbar^2}$ we get

$$\left[\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{2mr^2}{\hbar^2} \left(V(r) - E \right) \right] + \frac{1}{Y} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} \right] = 0$$

Where we seperated the terms that only depend on r and the terms that depend on θ and φ.
For this to hold true for all r, θ, φ, it must be orange and blue terms are constant. Let's say that they equal $l(l+1)$ and $-l(l+1)$ respectively.
We first solve the angular part. If we separate the variables

$$Y(\theta, \varphi) = T(\theta) \Phi(\varphi)$$

and plug this into the angular equation, we multiply everything by $\sin^2 \theta$ which gives us

$$\frac{1}{T}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial T}{\partial\theta}\right) + l(l+1)\sin^2\theta\right] = -\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial d\varphi^2}$$

where both sides again only depend on one variable. This also means that they equal a constant which we call m^2.

THe part in Φ is a simply differential equation with solution

$$\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\varphi^2} = -m^2 \implies \Phi(\varphi) = e^{im\varphi}$$

But since the exponential is $2\pi i$ periodic, it means that m must be an integer.
If we look at the equation for $T(\theta)$:

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial T}{\partial\theta}\right) + \left[l(l+1)\sin^2\theta - m^2\right] = 0$$

we can show that the solutions are given by the **associated Legendre functions** $P_l^m(x)$, which are defined as

$$P_l^m(x) := (-1)^m(1-x^2)^{\frac{m}{2}}\left(\frac{\partial}{\partial x}\right)^m P_l(x)$$

$$P_l^{-m}(x) := (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(x)$$

, where $P_l(x)$ is called the l-th **Legendre Polynomal**

$$P_l(x) := \frac{1}{2^l l!}\left(\frac{\partial}{\partial x}\right)^l (x^2-1)^l$$

Note that the solutions only make sense if l is an integer.
Also if $|m| > l$, then since $P_l(x)$ is a l-th order Polynomial $P_l^m(x) = 0$. The first couple associated polynomials look like this

$$P_0^0 = 1, \quad P_1^0 = \cos\theta, \quad P_1^1 = -\sin\theta$$

$$P_2^0 = \frac{1}{2}(3\cos^2\theta - 1), \quad P_2^1 = -3\sin\theta\cos\theta, \quad P_2^2 = 3\sin^2\theta$$

Combining these with $\Phi(\varphi)$, we get the angular wavefunctions, which are called **spherical harmonics** $Y_l^m(\theta, \varphi)$, which are given by

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}e^{im\varphi}P_l^m(\cos\theta)$$

which is defined for $l \in \mathbb{N}, m \in \{-l, \ldots l\}$. The first few spherical harmonics are

$$Y_0^0 = \sqrt{\frac{1}{4\pi}}$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos\theta, \quad Y_1^{\pm 1} = \mp\sqrt{\frac{3}{8\pi}} \sin\theta e^{\pm i\varphi}$$

$$Y_2^0 = \sqrt{\frac{5}{16\pi}}, \quad Y_2^{\pm 1} = \mp\sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta e^{\pm i\varphi}$$

$$Y_2^{\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2(\theta) e^{\pm 2i\varphi}$$

Note that the spherical harmonics are ortho-normal:

$$\int_0^\pi \int_0^{2\pi} \left(Y_{l'}^{m'}\right)^* Y_l^m \sin\theta d\theta d\varphi = \delta_{ll'}\delta_{mm'}$$

The differential equation for the radial part of the wavefunction is

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial R}{\partial r}\right) - \frac{2Mr^2}{\hbar^2}\left(V(r) - \Phi\right)R = l(l+1)R$$

And let $u(r) = rR(R)$, then

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial R}{\partial r}\right) = r\frac{\partial^2 u}{\partial r^2}$$

and we get the equivalent formulation

$$-\frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial r^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu$$

Which also simplifies the normalisation condition

$$\int_0^\infty r^2|R|^2 dr = 1 \iff \int_0^\infty |u|^2 dr = 1$$

3.9   The Hydrogen Atom

When we look at the spectrum of frequencies of the light emitted by atoms, we see that they always come in discrete frequencies instead of a whole range of them.

Hydrogen gas for example, when excited by a high voltage in a gas discharge tube has peaks at around $656.3\text{nm}, 486.1\text{nm}, 434\text{nm}$ and 410.1nm, which is called the Balmer series.

Johann Balmer, without explaining the wavelengths came up with an empirical formula for these wavelengths

$$\nu = \frac{c}{\lambda} \propto cR_\infty\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)$$

,where $R_\infty \simeq 1.1 \cdot 10^7\,\text{m}^{-1}$ is the Rydberg constant.

Before quantum mechanics, Niels Bohr constructed a model in which the electrons of the atom sit in stable orbits around the nucleus. The Bohr model however doesn't explain why the electrons emit radiation.

3.10   Radial equation for the Coulomb potential

We consider a hydrogen atom consisting of a proton and an electron. Since the mass of the proton is much bigger, we assume that it is motionless so we can center the coordinate system at the proton. The electron experiences a Coulomb potential given by

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0}\frac{1}{r}$$

If we insert this in to the radial equation from before, we get

$$-\frac{\hbar^2}{2m_e}\frac{\partial^2 u}{\partial r^2} + \left[\frac{e^2}{4\pi\varepsilon_0}\frac{1}{r} + \frac{\hbar^2}{2m_e}\frac{l(l+1)}{r^2}\right]u = Eu$$

Remember from the finite square well problem, that in order for the wave function to have a bound state (i.e. discrete Eigenvalues), the energy of the Eigenstate has to be less than the maximum value of the potential.
Here the potential is negative and approaches 0, for $r \to \infty$. Since we assume that the electron cannot "escape" the hydrogen atom, we assume that it's total energy is also less than zero. So we can focus on Eigenstates with $E < 0$.
So now let's define

$$\kappa := \frac{\sqrt{-2m_e E}}{\hbar} \in \mathbb{R}$$

We can then rewrite the radial equation as

$$\frac{1}{\kappa^2}\frac{\partial^2 u}{\partial r^2} = \left[1 - \frac{m_e e^2}{2\pi\varepsilon_0\hbar^2\kappa}\frac{1}{\kappa r} + \frac{l(l+1)}{(\kappa r)^2}\right]u$$

To make the equation prettier, we remove the dimensions by introducing the variables

$$\rho := \kappa r \quad \text{and} \quad \rho_0 := \frac{m_e e^2}{2\pi\varepsilon_0\hbar^2\kappa}$$

which allows us to rewrite the differential equation as

$$\frac{\partial^2 u}{\partial \rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2}\right]u$$

To further simplify the problem, we look at the asymptotic behavour $\rho \to \infty$ and $\rho to 0$. So

$$\frac{\partial^2 u}{\partial \rho^2} = u \quad \text{for} \quad \rho \to \infty$$

$$\frac{\partial^2 u}{\partial \rho^2} = \frac{l(l+1)}{\rho^2}u \quad \text{for} \quad \rho \to 0$$

because in each case, one term will dominate the other two.

For $\rho \to \infty$, the general solution will be of the form

$$u(\rho) = Ae^{-\rho} + Be^{\rho}$$

, for $u(\rho)$ to be finite, the Be^{ρ} must be zero.
For $\rho \to 0$, the general solution has the form

$$u(\rho) = C\rho^{l+1} + D\rho^{-l}$$

but because $l = 0, 1, 2, \ldots$ is positive, D must be zero in order for $u(\rho)$ to be finite for $\rho \to 0$.
Knowing the asymptotic behaviour, we can write the solution of the full differential equation as a product of the asymptotic solutions multiplied by a new function $v(\rho)$:

$$u(\rho) = e^{-\rho}\rho^{l+1}v(\rho)$$

We also know that ρ^{l+1} is the term with the smallest order, or else it would blow up for $\rho \to 0$, so we can write $v(\rho)$ as a polynomial of the form

$$v(\rho) = \sum_{j=0}^{\infty} c_j \rho^j$$

We then calculate the partial derivatives:

$$\frac{\partial u}{\partial \rho} = \rho^l e^{-\rho}\left[(l+1-\rho)v + \rho\frac{\partial v}{\partial \rho}\right]$$

$$\frac{\partial^2 v}{\partial \rho^2} = \rho^l e^{-\rho}\left(\left[-2(l+1) + \rho + \frac{l(l+1)}{\rho}\right]v + 2(l+1-\rho)\frac{\partial v}{\partial \rho} + \rho\frac{\partial^2 v}{\partial \rho^2}\right)$$

which when plugged into the full differential equation yields

$$\rho\frac{\partial^2 v}{\partial \rho^2} + 2(l+1-\rho)\frac{\partial v}{\partial \rho} + [\rho_0 - 2(l+1)]v = 0$$

Using the polynomial representation of $v(\rho)$

$$\frac{\partial v}{\partial \rho} = \sum_{j=0}^{\infty}(j+1)c_{j+1}\rho^j \quad \text{and} \quad \frac{\partial^2 v}{\partial \rho^2} = \sum_{j=0}^{\infty} j(j+1)c_{j+1}\rho^{j-1}$$

the differential equation becomes

$$\sum_{j=0}^{\infty} j(j+1)c_{j+1}\rho^j + 2(l+1)\sum_{j=0}^{\infty} j(j+1)c_{j+1}\rho^j - 2\sum_{j=0}^{\infty} jc_j\rho^j + [\rho_0 - 2(l+1)]\sum_{j=0}^{\infty} c_j\rho^j = 0$$

which, when evaluating the coefficients for each ρ^j, gives us

$$j(j+1)c_{j+1} + 2(l+1)(j+1)c_{j+1} - 2jc_j + [\rho_0 - 2(l+1)]c_j = 0$$

This gives us a recursive formula for the coefficents:

$$c_{j+1} = \left[\frac{2(j + l + 1) - \rho_0}{(j + 1)(j + 2l + 2)} \right] c_j$$

This looks like an infinite series, but it can't be. For $u(\rho) = 0$ as $\rho \to \infty$, we must have that the $e^{-\rho}$ domiates the polynomials.
This can only be true if the series terminates at some $j = N$ with $c_N = 0$ and $c_m = 0 \forall m \geq N$. The recursion formula then says

$$2(N + l) - \rho_0 = 0$$

, so if we define $n := N + l \in \mathbb{N}$ we get $\rho_0 = 2n$.
Therefore we can trace back our definitions

$$\rho_0 = \frac{m_e e^2}{2\pi\varepsilon_0 \hbar^2 \kappa} \quad \text{and} \quad \kappa = \frac{\sqrt{-2m_e E}}{\hbar}$$

to find that

$$E_n = -\frac{\hbar^2 \kappa_n^2}{2m} = -\left[\frac{m_e}{2\hbar^2} \left(\frac{e^2}{4\pi\varepsilon_0} \right)^2 \right] \frac{1}{n^2} =: \frac{E_1}{n^2} \quad \text{for} \quad n = 1, 2, 3, \ldots$$

We just were able to find out all the eigenenergies without even writing the wavefunctions. This shouldn't really be surprising, as in linear algebra we calculate the eigenvalues before the eigenvectors by finding the characteristic polynomial.
In some cases, knowing the eigenenergies is enough as we will see when we talk about spin for example.
But for the coulomb potential, we want to actually find the wave equations.
Before doing so, we can find the characteristic length scale for the wavefunctions.

$$a := \frac{\kappa}{n} = \frac{4\pi\varepsilon_0 \hbar^2}{m_e e^2} \simeq 0.53 \cdot 10^{-10} \, \text{m}$$

the **Bohr radius**.
To find the wave equations, recall that we split it into a radial and a spherical part.

$$\Psi_{nlm}(r, \theta, \rho) = R_{nl}(r) \cdot Y_l^m(\theta, \varphi)$$

which are labeled by what we call the **quantum numbers** n, l, m for

$$n \in \mathbb{N}, \quad l \in \{0, 1, 2, \ldots, n - 1\}, \quad m \in \{-l, -l + 1, \ldots, 0, \ldots, l - 1, l\}$$

We will see later that they are associated to the angular momentum. Also notice how the energy only depends on n. In fact the number of degenerate eigenstates with the same energy is

$$d(n) = \sum_{l=0}^{n-1} (2l + 1) = n^2$$

The ground state energy E_1 is

$$E_1 = -\left[\frac{m_e}{2\hbar b^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\right] = -13.6\text{eV}$$

And the only wave function with this energy is

$$\Psi_{100}(r,\theta,\varphi) = R_{10}(r)Y_0^0(\theta,\varphi)$$

Recall the recursion relation for the coefficients c_j from before:

$$c_{j+1} = \frac{2(j+l+1-n)}{(j+1)(j+2l+2)}c_j$$

so for $n = 1, j = 0$, this gives us $c_1 = 0$ and with $R = \frac{u}{r}$ and $\rho = \kappa r = \frac{r}{a}$ we get

$$u_{10} = c_0\rho e^{-\rho} \implies R_{10}(r) = \frac{c_0}{a}e^{-\frac{r}{a}}$$

Applying the normalisation condition, we can find out c_0 to be

$$\int_0^\infty |R_{10}|^2 r^2 dr \implies c_0 = \frac{2}{\sqrt{a}}$$

And combining it with $Y_0^0 = \frac{1}{\sqrt{4\pi}}$ which we found in the last lecture we get that

$$\Psi_{100}(r,\theta,\varphi) = \frac{1}{\sqrt{\pi a^3}}e^{-\frac{r}{a}}$$

We can now do this for all n, m, l. Let's see for example the wavefunctions with quantum number $n = 2$.

The radial wavefunction R_{20} is given by

$$R_{20} = \frac{c_0}{2a}(1 - \frac{r}{2a})e^{-\frac{r}{2a}}$$

and the states $|\Psi_{211}\rangle, |\Psi_{210}\rangle$ and $|\psi_{21-1}\rangle$ have the radial wavefunction

$$R_{21} = \frac{c_0}{4a^2}re^{-\frac{r}{2a}}$$

Note that the constants c_0 are not the same for each state and have to be determined by the normalisation condition.

More generally, the radial wavefunctions can be found by calculating the so called **associated Legendre Polynomials**

$$L_q^p(x) = \frac{x^{-p}e^x}{q!}\left(\frac{d}{dx}\right)^q (e^{-x}e^{p+q})$$

and by setting

$$U(\rho) = L_{n-l-1}^{2l+1}(2\rho)$$

Bringing everything together, the wavefunction (including normalisation) is given by

$$\Psi_{nlm}(r,\theta,\varphi) = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}}e^{-\frac{r}{na}}\left(\frac{2r}{na}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right)Y_l^m(\theta,\varphi)$$

Note that $v(\rho)$ is a polynomial or order $N - 1 = n - l - 1$, so the *radial* wavefunction has this many roots, and the *angular* wavefunction has $l - m$ roots as θ goes from 0 to π.

3.11   Angular momentum

In classical mechanics, we are often interested in finding conserved quantities and symmetries as they can greatly help us when analyzing problems.

In quantum mechanics, the conserved quantities come in the form of measurements of observables that commute with the Hamiltonian when the system is in an eigenstate of both the eigenstate and the Hamiltonian. In other words: we can measure the eigenvalues of eigenstates of operators that commute with the Hamiltonian.

A trivial operator would the Hamiltonian itself, so energy would be a conserved quantity.

Here we will define the **angular momentum** Operator(s), using the position operator $\vec{x}$ and momemtum Operator $\vec{p}$.

$$\vec{\hat{L}} = \vec{\hat{x}} \times \vec{\hat{p}}$$

, or for each of the three components:

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \quad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z, \quad \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$$

This definition follows the classical definition of angular momentum $\vec{L} = \vec{r} \times \vec{p}$.

we will from now on stop writing the hats as it should be clear from context.

Let's consider the commutator between two of the components. We see that they do not commute.

$$\begin{aligned}
[L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\
&= [yp_z, zp_x] - \underbrace{[yp_z - xp_z]}_{=0} - \underbrace{[zp_y, zp_x]}_{=0} + [zp_y, xp_z] \\
&= yp_x[p_z, z] + xp_y[z, p_z] = i\hbar(xp_y - yp_x) \\
&= i\hbar L_z
\end{aligned}$$

This means that the different components of angular momentum are simultaneously observable, so it doesn't really make sense to think of angular momentum as a vector as it's direction is fuzzy. Similarily, we get the other commutators

$$[L_y, L_z] = i\hbar L_x, \quad \text{and} \quad [L_z, L_x] = i\hbar L_y$$

Note that the square $L^2 = L \circ L$ does commute with each of the components

$$\begin{aligned}
[L^2, L_x] &= [L_x^2, L_x] + [L_y^2, L_x] + [L_z^2, L_x] \\
&= L_y[L_y, L_x] + [L_y, L_x]L_y + L_z[L_z, L_x] + [L_z, L_x]L_z \\
&= L_y(-i\hbar L_z) + (-i\hbar L_z)L_y + L_z(i\hbar L_y) + (i\hbar L_y)L_z = 0
\end{aligned}$$

similarly: $[L^2, L_y] = [L^2, L_z] = 0$

This means that sould be able to find simultaneous eigenstates of L^2 and L_z. Let's denote the eigenstates as $|l, m\rangle$, so the equaion will look like

$$L^2 |l, m\rangle = \hbar^2 l(l+) |l, m\rangle \quad \text{and} \quad L_z |l, m\rangle = \hbar m |l, m\rangle$$

where the $\hbar$ are such that l and m are dimensionless.

Now recall the ladder operators a and $a^\dagger$ from the quantum harmonic oscillator, with which were able to create new eigenstates for our system. They satisfied the commutator relations

$$[a^\dagger a, a^\dagger] = a^\dagger \quad \text{and} \quad [a^\dagger a, a] = -a$$

We can find a similar relation between the operator pairs

$$L_\pm := L_x \pm iL_y$$

which gives the following commutator relations

$$[L_z, L_\pm] = \pm\hbar L_\pm, \quad [L^2, L_\pm] = 0$$

where we can think of L_z as being like $a^\dagger a$ and $L_\pm$ as $a^\dagger$ and a.

So we can create new states like this:

$$L^2(L_\pm |l, m\rangle) = L_\pm L^2 |l, m\rangle = \hbar^2 l(l+1)(L_\pm |l, m\rangle)$$
$$L_z(L_\pm |l, m\rangle) = (\pm\hbar L_\pm + L_\pm L_z) |l, m\rangle = \hbar(m \pm 1)(L_\pm |l, j\rangle)$$

which shows that $L_\pm |l, m\rangle$ is also an eigenstate of L^2 with eigenvalue $\hbar^2 l(l+1)$ and that $L_\pm |l, m\rangle$ is an eigenstate of L_z with eignevalue $\hbar(m \pm 1)$ We can bring this into a nicer form by setting

$$A^m_{l,\pm} := \hbar\sqrt{l(l+1) - m(m \pm 1)}$$

to get the ladder relation similar to what we found in the harmonic oscillator

$$L_\pm |l, m\rangle = A^m_{l,\pm} |l, m \pm 1\rangle$$

likewise, we can again show what kinds of values l and m can take by considering the following:
The operator L_+ lets us increase the eigenvalue of L_z by $\hbar$ and L_- decreases the eigenvalue. But this process has to end somewhere, or else we would have measured that the the z-component of the angular momentum has a bigger magnitude that the total angular momentum itself.
So there must be some maximal value of m, say $\mu := m_{\max}$ aswell als a minimal value $\mu' := m_{\min}$ such that the Ladder operators become zero:

$$L_+ |l, \mu\rangle = 0 \quad \text{and} \quad L_- |l, \mu'\rangle = 0$$

we are also able to find out these relations

$$L_\pm L_\mp = (L_x \pm iL_y)(L_x \mp iL_y) = L_x^2 + L_y^2 \mp i(L_x L_y - L_y L_x) = L^2 - L_z^2 \pm \hbar L_z$$

solving the above for L^2, we see that

$$L^2 |l, \mu\rangle = (L_- L_+ + L_z^2 + \hbar L_z) |l, \mu\rangle = \hbar^2 \mu(\mu + 1) |l, \mu\rangle$$

but we also knew that L^2 fulfilled the eigenstate equation

$$L^2 |l, \mu\rangle = \hbar^2 l(l+1) |l, \mu\rangle$$

so we immediately see that $m_{\max} = \mu = l$. In a similar fashion we can show

$$L^2 \left|l, \mu'\right\rangle = \hbar^2 \mu'(\mu' - 1) \left|l, \mu'\right\rangle$$

This equation on itself has two solutions, $\mu' = -l$ and $\mu' = l + 1$, but the latter would mean $\mu' > \mu$. To see what values l can take on we notice that the interval $[-l, l]$ must have the length of an integer because we take an integer multiple times the application of L_+ to increase m by one starting from $-l$ to l.

So the possible values for l, and m are

$$l = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \ldots \quad \text{and} \quad m = -l, -l + 1, \ldots l + 1, l$$

which look suspiciously like the possibles values of l and m for the spherical harmonics except that now we can have half-integer values. Also notice that the maximum measured result for L_z, given by

$$L_z \left|l, m\right\rangle = \hbar m \left|l, m\right\rangle \implies L_{z,\max} = \hbar l$$

from before is always lass than the total magnitude $\hbar\sqrt{l(l+1)}$ (except for $l = 0$) which we got from

$$L^2 \left|l, \mu\right\rangle = \hbar^2 \mu(\mu + 1) \left|l, \mu\right\rangle$$

To further strenghten the connection between the eigenstates $\left|l, m\right\rangle$ and the spherical harmonics Y_l^m, let's re-write the angular momentum operators in spherical coordinates. In the position representation, it is:

$$\vec{L} = -i\hbar(\vec{r} \times \vec{\nabla})$$

so using the spherical nabla operator

$$\vec{\nabla} = \vec{r}\frac{\partial}{\partial r} + \vec{\theta}\frac{1}{r}\frac{\partial}{\partial r} + \vec{\varphi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}$$

we can find that the cartesian components L_x, L_y, L_z of the angular momentum are given by

$$L_x = i\hbar\left(\sin\varphi\frac{\partial}{\partial\theta} + \cos\varphi\cot\theta\frac{\partial}{\partial\varphi}\right)$$

$$L_y = i\hbar\left(-\cos\varphi\frac{\partial}{\partial\theta} + \sin\varphi\cot\theta\frac{\partial}{\partial\varphi}\right)$$

$$L_z = -i\hbar\frac{\partial}{\partial\varphi}$$

aswell as the square of the Angular momentum operator

$$L^2 = -\hbar\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\frac{\partial^2}{\partial\varphi^2}}\right]$$

now recall from spherical harmonics that our equation for the angular part was

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}Y_{lm}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}Y_{lm}\right] = -l(l+1)Y_{lm}$$

which is equivalent to the eigenvalue equation

$$L^2\left|l, m\right\rangle = \hbar^2 l(l+1)\left|l, m\right\rangle$$

and likewise for the phi component, the equation

$$\frac{\partial^2}{\partial\varphi^2 Y_{lm}} = -m^2 Y_{lm}$$

is equivalent to the eigenvalue equation

$$L_z\left|l, m\right\rangle = \hbar m\left|l, m\right\rangle$$

Also, the Hamiltonian can be written as

$$H = \frac{1}{2mr^2}\left[-\hbar^2\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + L^2\right] + V$$

so any eigenstate of the central potential hamiltonian H is simultanously an eigenstate of L^2 and L_z

$$H\psi = E\psi, \quad L^2\psi = \hbar^2 l(l+1)\psi, \quad L_z\psi = \hbar m\psi$$

3.12   Spin

The connection between angular momentum and spherical harmonics only works for the integer values of l, so we still need to discuss the half-integer values of l.

Although half-integer values might seem odd, there are physical experiments that point toward the existence of a quanttity that behaves like angular momentum aswell as other experimental observations that hint at the existence of half-integer values of l.

In particular, when taking into account special relativity in a high-resolution spectroscopy of hydrogen, we see two spectral lines for the Balmer series transition $l = 1 \rightarrow l = 0$, where we would only expect $2l + 1$ of them, which is an odd number for integer values of l.

Another example that we will look at is an experiment performed by Otto Stern and Walther Gerlach in 1922, in which they set out to measure the magnetic moment of atoms. Recall that the magnetic moment $\vec{\mu}$ of a current I in a loop with radius r is given by

$$\vec{\mu} = \pi r^2 I\hat{n}$$

, where $\hat{n}$ is the unit normal vector perpendicular to the plane in counter-clockwise direction. In the Bohr model of the atom, the electron has a circular orbit around the nucleus.

So if v is the speed of the electron, e its charge and r the radius of the orbit, then the current I is

$$I = -e\left(\frac{v}{2\pi r}\right) \implies \vec{\mu} = -\pi r^2 e \frac{v}{2\pi r}\hat{n} = -\frac{e}{2m}\vec{r} \times \vec{p} = -\frac{e}{2m}\vec{L}$$

so when the atom is placed in a magnetic field $\vec{B}$ along the z direction, it would experience a force $\vec{F}$ equal to the gradient of the Potential $-\vec{\mu} \cdot \vec{B}$

$$\vec{F} = -\nabla(-\vec{\mu} \cdot \vec{B}) = \mu_z \frac{\partial B_z}{\partial z}\hat{z} \propto L_z \frac{\partial B_z}{\partial z}\hat{z}$$

Here we assume that the field gradient are mostly along the z direction, so atoms with different z component of angular momentum will experience different forces. But also, we know that L_z can only take on $2l + 2$ values which should be an odd number for integer l

The Stern-Gerlach experiment used an apparatus, where silver atoms pass through an inhomogeneous filed and land on a glass plate with the location where tey land indicating the force that the atoms experienced from the magnetic field.

They observed that not only were the measured outcomes quantized, they found that there were two distinct locations.

To explain this, a new property called **spin** was proposed which behaved similar to angular momentum in that the **spin operator** S satisfied the commutator and eigenvalue equations

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y$$
$$S^2 |s, m_s\rangle = \hbar^2 s(s+1) |s, m_s\rangle \quad \text{and} \quad S_z |s, m_s\rangle = \hbar m |s, m_s\rangle$$

where s and m_s can take on the values

$$s = 0, \tfrac{1}{2}, 1, \frac{3}{2}, \dots \quad \text{and} \quad m_s = -s, -s+1, \dots, s-1, s$$

It turns out that unlike angular momentum, each elementary particle has a specific value of s that is conserved, i.e. never changes. Electrons have spin $\frac{1}{2}$, photons have spin 1 and so on.

The most important case is $s = \frac{1}{2}$. Most particles that make up ordinary matter are leptons and have spin $\frac{1}{2}$. This also means that they can have only two eigenstates $m = -\frac{1}{2}, \frac{1}{2}$ which is nice because qunatum systems that only have two eigenstates can be described by spin-$\frac{1}{2}$ particles, even if the system has no particles with such spin. It also features a quantum analogue of the classical bit with two states 0 and 1, which in the quantum world are called qubits.

We often describe spin-$\frac{1}{2}$ particles in vector-notation using the eigenstates as a basis:

$$|l = \tfrac{1}{2}, m_s = -\tfrac{1}{2}\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |l = \tfrac{1}{2}, m_s = \tfrac{1}{2}\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

where a general state $|\psi\rangle$ can be written as

$$|\psi\rangle = \alpha |\tfrac{1}{2}, \tfrac{1}{2}\rangle + \beta |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \leftrightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

We've seen how operators acting on states can be viewed as matrices. In our case, we have a 2-dimensional vector space so for example, we can write

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} =: \frac{\hbar}{2}\sigma_z$$

we see that the eigenstates to $s = \frac{1}{2}$ are indeed eigenvectors are indeed eigenvectors of the matrix σ_z/eigenstates of the $\hat{S}_z$ operator

$$\hat{S}_z |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\hat{S}_z |\tfrac{1}{2}, \tfrac{1}{2}\rangle = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Recall from linear algebra if we have an operator $\hat{O}$ and some orthonormal basis $\{|\varphi_n\rangle\}$, then we can recover the matrix coefficents o_{ij} as follows

$$o_{ij} = \langle \varphi_i|\hat{O}|\varphi_j\rangle$$

so if we want to find out what the matrix σ_+ representing the S_+ operator looks like, we can calculate $\langle \frac{1}{2}, m|S_+|\frac{1}{2}, m'\rangle$ for $m, m' = -\frac{1}{2}, \frac{1}{2}$.
Because the ladder operator satisfies

$$S_\pm |s, m\rangle = \hbar\sqrt{s(s+1) - m(m\pm 1)}\,|s, m\pm 1\rangle$$

we know that

$$\langle \tfrac{1}{2}, \tfrac{1}{2}|\hat{S}_+|\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \hbar\,\langle \tfrac{1}{2}, \tfrac{1}{2}\rangle = \hbar$$

and it vanishes for all the others. Therefore, the matrix representation σ_+ is given by

$$\hat{S}_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \implies \quad \begin{pmatrix} 1 & 0 \end{pmatrix}\cdot\hbar\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \hbar\checkmark$$

Similarly, for S_- we find

$$S_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and using the relations $S_x = \frac{1}{2}(S_+ + S_-), S_y = \frac{1}{2i}(S_+ - S_-)$, we get

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} =: \frac{\hbar}{2}\sigma_x \quad \text{and} \quad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} =: \frac{\hbar}{2}\sigma_y$$

We call $\sigma_x, \sigma_y, \sigma_z$ the **Pauli matrices**, which are just the matrix representations of the corresponding Operators.

For example, if we ask what the probability of a state $|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$ measuring as an eigenvector of S_x is, we take the dot product with the normed eigenstate

$$|\langle \varphi | \psi \rangle|^2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \frac{|a + b|^2}{2}$$

Now we can explain the Stern-Gerlach experiment by particles which have spin $\frac{1}{2}$. The two eigenstates with this spin experience a force in opposite directions.
By doing this experiment with electrons, we can determine the magnetic moment of the electrons, which was found to be

$$\mu = \pm \frac{e\hbar}{2m_e} =: \pm \mu_B$$

where μ_B is the called the **Bohr magneton**