MPhys Notes

Quantum Physics and Relativity

Dave Scott

Based on a course by Prof. Brian Cox and Prof. Jeff Forshaw

July 4, 2018

Contents

1 Relativity **3**
1.1 Lexical Preface . . . 4
1.2 Spacetime and Four-Vectors . . . 5
1.3 Galilean Relativity . . . 6
1.4 The Principles of Einstein's Special Theory of Relativity . . . 8
1.5 Time Dilation and Length Contraction . . . 8
1.5.1 Time Dilation . . . 8
1.5.2 Length Contraction . . . 10
1.5.3 Evidence . . . 11
1.6 Proper Time and Lorentz Transformations . . . 13
1.7 Velocity Transformations and the Doppler Effect . . . 15
1.7.1 Four-Velocity . . . 15
1.7.2 Velocity Transformations . . . 16
1.7.3 The Doppler Effect . . . 17
1.8 Causality . . . 17
1.9 Energy and Momentum . . . 18

2 Quantum Mechanics **19**
2.1 History . . . 19
2.2 Units . . . 21
2.3 Matter Waves, Wavefunctions and Quantum States . . . 21
2.4 The de Broglie Relation . . . 23
2.5 Heisenberg's Uncertainty Principle . . . 24
2.6 Applications of Quantum Mechanics . . . 27
2.7 Observables and Operators . . . 27
2.8 The Time-Independent Schrödinger Equation (TISE) . . . 28
2.9 The Momentum Operator . . . 29

Chapter 1

Relativity

[I]t is impossible to begin to learn that which he thinks he already knows.

— Epictetus*, Discourses, Book II. Ch. 17*

Time is relative. Start with this premise and this course will make far more sense than attempting to disprove the opposite.

Rather confusingly, relativity is pretty much the only university course in which incorrect notions are taught prior to correct notions. For instance, we do not pre-teach Ptolemic refraction in optics, we do not pre-teach caloric theory in thermodynamics and we do not pre-teach contact tension in electromagnetism. Whilst intended for people of all skill levels, my notes are predominantly targetted toward university students and so I will follow a typical university course structure and include the incorrect introduction, though personally I would prefer not to.

1.1 Lexical Preface

Definition 1.1: Relativity

In physics, **relativity** refers to the dependence of various physical phenomena on the motion of an **observer**. More generally, relativity depends on the viewpoint of an observer (subjective), in contraposition to absoluteness which does not (objective).

Definition 1.2: Observers

In general, an **observer** is a person or machine capable of taking measurements. For most examples, one can think of an observer as a literal person watching an **event**. The position in which the observer resides is known as a **frame of reference**.

Definition 1.3: Events

An **event** is the result of some physical phenomenon, e.g. collision of two spheres, decay of a radioactive nucleus, passing of a train.

Definition 1.4: Reference Frames

A **reference frame** or **frame of reference** is the means by which events can be described in space and time. All frames move relative to one other, with those travelling at constant relative speeds being known as **inertial** frames. For the special case in which the relative velocity is zero, the frames are said to reside in the **rest frame**.

Definition 1.5: Invariance

The term **invariance** means that the properties of a system remain unchanged under some **transformation**, e.g. rotation, translation.

Definition 1.6: Transformations

A **transformation** is the mathematical relation of properties between systems, most commonly that of positional coordinates.

1.2 Spacetime and Four-Vectors

The story goes that as René Descartes lay in his bed one day, he watched a fly walking around on his ceiling. He asked himself how he might describe its location and concluded that he could do so by measuring its distance from each wall. However, the distance from one wall is encompassed in the distance from the wall opposite, hence with only two numbers, the fly's location could be denoted. The problem was that once the fly took flight, this method was no longer a valid description of its location as the two numbers could no longer uniquely identify the fly's position. To resolve this, a third number was introduced - measured as the distance from either the ceiling or the floor - to denote the height of the fly. These three numbers were simply labelled x, y and z and with these any position in space could be described mathematically, and thus the Cartesian coordinate system came to be. A shorter way of denoting the position of an object is by writing the components along each dimension in a container known as a **three-vector** - so called because it is a vector with three components, e.g. the position vector, $\mathbf{x}$, expands to $\mathbf{x} = (x, y, z)$.

The problem with the Cartesian coordinate system is that it is limited to only being able to describe the spatial location of an object at a given moment in time. Therefore, a further generalisation to this coordinate system would be to specify at what point in time the object existed at a point in space and hence the concept of time as a dimension is born.

The ability to specify the position of an object in both spatial and temporal coordinates was a breakthrough, however the question as to what these four dimensions were actually measuring remained. In 1908, German mathematician Hermann Minkowski found that space and time could be interwoven into a singular four-dimensional "space", known as **spacetime** or **Minkowski space**. It is from this that it became possible to add a fourth component to the position vector, with it thus becoming a **four-vector**. The position four-vector is denoted x, i.e. $x = (t, x, y, z)$, however one should be careful not to confuse this with the x-axis.

1.3 Galilean Relativity

Galilean relativity describes the invariant nature of matter and time between frames of reference in Newtonian mechanics. As with the underpinnings of classical mechanics, the foundations of this relativity are laid down as axioms based upon intuition.

Relativity states that there is no preferred inertial frame as the dynamics should be equivalent between observers. Consider two frames of reference, O and O', moving with a non-zero relative velocity. There are three ways of viewing this:

- O is stationary, O' is moving
- O is moving, O' is stationary
- Both O and O' are moving relative to some outside observer

The foundations of relativity are built upon the fact that all of these are the same and that it simply depends on the observer as to how one goes about describing the system.

Example 1.1

Examples used in relativity courses are nearly always limited to trains, particles or spaceships; these are useful because they generally have few external effects associated with them, e.g. friction. Furthermore, for the purpose of this course all frames are considered to be inertial.

To be in keeping with convention, consider two people sat on different trains which are on parallel, flat, rigid, frictionless tracks. Neither of them undergo acceleration or deceleration - hence inertial - and the outside cannot be seen except for through one window and then only when the other train travels past. The question is: what measurement could be made to determine which of them is moving?

Crafty readers may reply that they would try to look at something other than the train opposite, such as the landscape/sky or potentially look out of a different window on the other side/front/back of the train. To these people I say simply that this is not a riddle!

More honest but slightly naïve readers may offer a number of suggestions such as shining a light or throwing a ball across the gap and measuring which way it gets deflected. The issue with these experiments is that they can only measure the velocity of the other train relative to the observer - not which is actually in motion.

The answer is that there is no experiment to prove which train is in motion, thus only the motion with respect to one another is all that is necessary.

Consider an observer on a train at position $\mathbf{x} = (x, y, z)$ moving at constant velocity v along the x-axis, with a second observer stood stationary on a platform located at $\mathbf{x}' = (x', y', z')$. Each observer has one of two identical clocks, with the time observed on the train being denoted t while the time measured on the platform is denoted t'. At an initial time $t = t' = 0$ the observers are considered to have the same origin; $\mathbf{x} = \mathbf{x}'$, i.e. $x = x'; y = y'; z = z'$. The relation between the positions of the observers in classical mechanics is known as a **Galilean transformation**. In Newtonian mechanics the time measured on the train is equal to the time measured on the platform, hence $t' = t$. The locations of the observers can be calculated via Newton's laws of motion, resulting in the conclusion that the train has moved a distance vt away from the platform along the x-axis with no alteration in the y- or z-axes. To briefly summarise, these can be written as

$$
\begin{aligned}
t' &= t \\
x' &= x - vt \\
y' &= y \\
z' &= z.
\end{aligned}
$$

This is a Galilean transformation.

According to Newtonian mechanics speeds are additive, i.e. if the observer on the train throws a ball forward at a speed v' then the observer on the platform will measure the ball as moving with a speed of $v + v'$. Herein lies a major issue, but do not expect to understand why this is an problem just yet.

1.4 The Principles of Einstein's Special Theory of Relativity

The downfall of Galilean relativity is that it assumes measurements of time are equal in all frames of reference. This issue derives from Newtonian mechanics which has a number of other downfalls, such as the additive nature of velocities and the absoluteness of space.

Einstein's special theory of relativity was initially built upon two postulates:

1. All physical laws are the same (invariant) in all inertial frames
2. The speed of light in a vacuum is equal for all inertial frames of reference

These postulates are only applicable to inertial frames, i.e. frames with constant velocity, and it is for this reason that acceleration and deceleration are ignored for the purpose of this course.

1.5 Time Dilation and Length Contraction

1.5.1 Time Dilation

Consider two observers in separate inertial frames with relative velocity v; one standing on a platform and the other on a train. One observer has two parallel mirrors with light bouncing between, known as a light clock. Figure 1.1 demonstrates how the paths appear to the different observers. For the observer in the rest frame, O, who measures time t, the distance travelled in half a clock cycle (one direction; without return) is

$$l = ct. \tag{1.1}$$

The (squared) distance travelled according to the second observer, in frame O', is given by Pythagoras' theorem as

$$(ct')^2 = l^2 + (vt')^2. \tag{1.2}$$

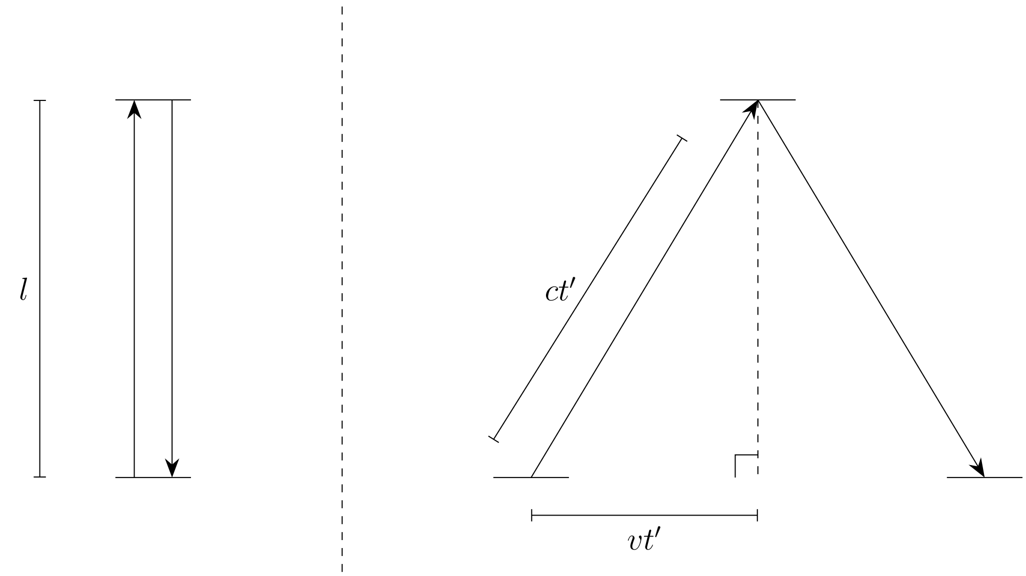

Figure 1.1: Left: View of light clock in rest frame.
Right: View of light clock from outside observer.

Substituting the mirror separation from Equation (1.1) into Equation (1.2) and rearranging, one gets

$$(ct')^2 = (ct)^2 + (vt')^2$$
$$t'^2 \left(1 - \left(\frac{v}{c}\right)^2\right) = t^2.$$

The fraction $\frac{v}{c}$ is sometimes called the relativistic relative velocity as it relates the velocity of the particle to that of light and is denoted β, hence

$$t' = \frac{1}{\sqrt{1-\beta^2}} t.$$

The factor relating the observed times betwen frames is denoted γ and is known as the **gamma** or **Lorentz factor**;

$$t' = \gamma t. \tag{1.3}$$

This is time dilation.

The gamma factor is limited to values between 1 and ∞, corresponding to speeds of $v = 0$ and $v = c$, respectively. The implication of this is that the observer in the "moving" frame O' measures time outside the train as travelling much faster than in their rest frame, whereas the "stationary" observer in frame O sees time on the train as travelling much slower than their own. The time experienced by a clock is known as the **proper time**, i.e. the time that has elapsed in the clock's rest frame, which is denoted τ, whereas the observed time changes notation to t, hence the equation for the proper time of an inertial frame is

$$t = \gamma\tau.$$

1.5.2 Length Contraction

According to the first postulate of relativity, within every frame of reference a speed is given by the ratio of distance covered to the time elapsed within the frame. Therefore two inertial frames measuring speeds within their own frames are given by

$$v = \frac{x}{t} \text{ and } v = \frac{x'}{t'}.$$

These can be related to one another by Equation (1.3) as

$$\begin{aligned} \frac{x}{t} &= \frac{x'}{t'} = \frac{x'}{\gamma t} \\ \Rightarrow x' &= \frac{1}{\gamma}x. \end{aligned}$$

This is length contraction.

The consequence of this is that the observer on the train "in motion" measures objects outside of the train as being shorter than when in the same frame of reference. Furthermore, the platform observer measures the train as being shorter when it is "in motion" versus when at rest.

Example 1.2

Consider a rod of length L_0 travelling at velocity $v = \frac{\sqrt{5}}{3}c$ relative to an observer. In the rod's rest frame it still measures itself as being L_0 in length, however the observer measures the rod as being of length $\frac{1}{\gamma}L_0$, i.e.

$$\begin{aligned} L &= \frac{1}{\gamma}L_0 \\ &= \sqrt{1 - \left(\frac{\sqrt{5}}{3}\right)^2} L_0 \\ &= \frac{2}{3}L_0. \end{aligned}$$

1.5.3 Evidence

Perhaps a few skeptical readers think that this is simply a matter of mathematics which do not pertain to the real world. In an attempt to win these people over - as well as further cementing the ideas for those more trusting readers - some empirical evidence will now be presented and explained.

Frisch-Smith Experiment

In 1963 two American physicists, David H. Frisch and James H. Smith, studied the flux and velocities of particles called muons from a scientific research field station on top of Mount Washington, New Hampshire, which is 1900 m above sea level. Muons are short-lived particles, with mean lifetimes of 2.2 µs, that are created by cosmic rays interacting with the upper atmosphere. After a time t the proportion of particles, $\frac{N}{N_0}$, with mean lifetime τ is given by

$$\frac{N}{N_0} = e^{-\frac{t}{\tau}}.$$

A second detector at sea level also measured the flux and velocities of muons with both experiments measured the velocities to be $0.99c$, where c is the speed of light in a vacuum. Classically it therefore takes 6.4 µs to traverse the 1900 m difference in altitude, hence one then expects the proportion of muons reaching the lower detector to be

$$\frac{N}{N_0} = e^{-\frac{t}{\tau}} = e^{-\frac{6.4\ \mu\text{s}}{2.2\ \mu\text{s}}} = 0.05 = 5\%.$$

However, this assumes that the time elapsed according to the muons is equal to that by an outside observer, which is not the case. The Lorentz factor of the muons is $\gamma = \frac{1}{\sqrt{1-0.99}} = 10$, hence the time observed by the muon is $\frac{1}{10} \times 6.4\ \mu\text{s} = 0.64\ \mu\text{s}$ and thus the muon survival ratio given by relativity is

$$e^{-\frac{t}{\tau}} = e^{-\frac{0.64\ \mu\text{s}}{2.2\ \mu\text{s}}} = 0.75 = 75\%.$$

Frisch and Smith observed an average flux of 563 muons per hour from the observatory detector and 412 muons per hour from the detector at sea level, resulting in a survival rate of just over 73%, a remarkably accurate prediction of special relativity.

GPS Navigation

Whilst special relativity only considers inertial frames, general relativity extends the framework to incorporate frames with acceleration. Whilst the derivation and deeper understanding of general relativity is far beyond this course, one of the main ideas is that gravitation and acceleration are equivalent, with both resulting in time dilation and length contraction.

The Global Positioning System (GPS) provides time and location information to the globe and is best known for its use in navigation systems, e.g. Google/Apple maps, in-car navigation devices. Since the information is transmitted via an array of satellites orbitting the Earth, effects relating to gravity must be accounted for. General relativity states that the highly accurate atomic clocks on the satellites will run slower than those on Earth by approximately 40 µs per day, as such the output must be slowed by this amount. If this adjustment is not made then the maps would become inaccurate by about 11 km per day, hence after one month the GPS maps would be inaccurate by around 330 km. That GPS systems are still accurate to within 10 m after years of running is a testament to the accurate predictions of general relativity.

1.6 Proper Time and Lorentz Transformations

Consider a single point located at the origin of a frame, $(0,0,0)$, from which a sphere of light is emitted. The radius, r, of the sphere expanding at the speed of light, c, at a time, t, is given by $r = ct$. However, according to Pythagoras' theorem, any point on the spherical surface is also given by $r^2 = x^2 + y^2 + z^2$. These radii must be equivalent, hence one can write

$$c^2t^2 - x^2 - y^2 - z^2 = 0. \tag{1.4}$$

One can rearrange the equation to incorporate proper time:

$$\begin{aligned} t &= \gamma\tau \\ \rightarrow \tau^2 &= \frac{1}{\gamma^2}t^2 \\ &= t^2\left(1 - \left(\frac{\mathbf{v}}{c}\right)^2\right) \\ &= t^2 - \left(\frac{\mathbf{x}}{c}\right)^2 \\ \Rightarrow c^2\tau^2 &= c^2t^2 - x^2 - y^2 - z^2. \end{aligned}$$

Furthermore, Equation (1.4) can be rearranged as

$$c = \frac{\sqrt{x^2 + y^2 + z^2}}{t} = \frac{|\mathbf{x}|}{t},$$

i.e. $c = v$, hence the proper time is zero for anything travelling at the speed of light. As such, anything travelling at the speed of light would be able to travel to any future point in time without experiencing any time itself. Positive values of proper time, i.e. $c^2t^2 - x^2 - y^2 - z^2 > 0$, correspond to a universe in which all motion is limited to the speed of light, such as our own. Negative values of proper time instead correspond to motion in which the speed of light is the *minimum* speed, however the time becomes imaginary and so loses any classical meaning.

Consider two inertial frames, O and O', with identical clocks, synchronised such that $t = t' = 0$ when the origins coincide, i.e. $x = x'; y = y'; z = z'$. At some later time, t, the rest frame emits a sphere of light and at another time after original coincidence, t', the moving frame emits its own sphere of light. At some point these spheres must coincide, hence the proper times for the rest frame, A, and moving frame, B, must be equal, i.e. $c^2\tau_O^2 = c^2\tau_{O'}^2$, so

$$A \equiv c^2t^2 - x^2 - y^2 - z^2 = 0$$
$$\text{when } B \equiv c^2t'^2 - x'^2 - y'^2 - z'^2 = 0.$$

As the proper times are quadratic with respect to their variables but must be equal at coincidence, they must be related linearly, i.e.

$$A = kB,$$

where k is some undetermined linear constant. Since the motion is considered to be along the x-axis only, x' must vary linearly with x and t, hence the transformation of the spatial components have the form

$$\begin{aligned} x' &= k\alpha x + k\beta t \\ y' &= ky \\ z' &= kz. \end{aligned}$$

Since any motion in either y or z directions is parallel between the frames, the components must be equal, hence $y = y'$ and $z = z'$, then $k = 1$, then $x' = \alpha x + \beta t$. At an initial time $t = t' = 0$, it is known that $x' = \gamma x$ and thus $\alpha = \gamma$. Considering the primed frame, the components along the x-axis are $x' = 0$ and $x = vt$, then

$$0 = \gamma vt + \beta t$$

and as such $\beta = -\gamma v$ so that now

$$x' = \gamma(x - vt). \tag{1.5}$$

By considering the unprimed frame with the same reasoning such that $x' = -vt'$ and $x = 0$ then

$$x = \gamma(x' + vt'). \tag{1.6}$$

Substituting Equation (1.5) into Equation (1.6) or vice versa, then rearranging, one gets the times in the frames as

$$\begin{aligned} t &= \gamma\left(t' + \frac{v}{c^2}x'\right) \\ t' &= \gamma\left(t - \frac{v}{c^2}x\right). \end{aligned}$$

In summary;

$$\begin{aligned} t &= \gamma\left(t' + \frac{v}{c^2}x'\right) & t' &= \gamma\left(t - \frac{v}{c^2}x\right) \\ x &= \gamma(x' + vt') & x' &= \gamma(x - vt) \\ y &= y' & y' &= y \\ z &= z' & z' &= z. \end{aligned}$$

These are Lorentz transformations.

1.7 Velocity Transformations and the Doppler Effect

1.7.1 Four-Velocity

The four-vector of position, x, is given by $x = (ct, x, y, z)$. As in classical mechanics, the four-vector of velocity is given by the time derivative of the position. However, since the rate of change of position is with respect to spacetime rather than simply spatial coordinates, the derivative of the four-vector is from its rest frame and so is with respect to proper time. The **four-velocity**, v, is thus given by

$$\begin{aligned} v \equiv \frac{dx}{d\tau} &= \frac{dt}{d\tau}\frac{dx}{dt} \\ &= \gamma\frac{dx}{dt} \\ &= \gamma\left(\frac{d(ct)}{dt}, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \\ &= \gamma(c, v_x, v_y, v_z), \end{aligned}$$

where $v_{x,y,z}$ are the velocities in the respective directions.

1.7.2 Velocity Transformations

Example 1.3

Consider a person driving along an autobahn at close to the speed of light because they have an awesome car and are allowed to do so. The driver notices that it is getting close to night time and so turns on the headlights, at what speed does the driver observe the headlight beams escaping them?

Since the speed of light, c, is a constant universal speed limit, an outside observer cannot see the headlight beams travelling faster (or slower) than the speed of light. According to Newtonian mechanics velocities are additive, and so the driver would observe the beams escaping at the difference of velocities, i.e. when the car is travelling at speed v the light escapes at velocity $c-v$. Since an outside observer would see the light escaping at c and the driver does not, this scenario is in conflict with both postulates of relativity. How then can this seemingly paradoxical scenario be resolved?

The means to answer this example comes in the form of a **velocity transformation**, that is, a means of describing motion in two inertial frames that are moving relative to one another with a speed v. The velocity of an object within frame O', u', is given by the time derivative, t', of the space traversed in the frame, x', i.e.

$$u' = \frac{dx'}{dt'}.$$

From the Lorentz transformation it is known that $dx' = \gamma(dx - vdt)$ and $dt' = \gamma\left(dt - \frac{vdx}{c^2}\right)$, hence

$$\begin{aligned} u' &= \frac{\gamma(dx - vdt)}{\gamma\left(dt - \frac{vdx}{c^2}\right)} \\ &= \frac{\frac{dx}{dt} - v}{1 - \frac{v\frac{dx}{dt}}{c^2}} \\ &= \frac{u - v}{1 - \frac{uv}{c^2}}. \end{aligned}$$

This therefore relates the speed at which an object is seen to move, u', in an inertial frame, O, which is moving with relative velocity v when the object is seen to move at a speed u within its frame of reference, O'.

1.7.3 The Doppler Effect

Consider a source of light with wavelength λ moving toward an observer at a speed v. The measured time per oscillation of the light, t, by the observer is given by

$$t = \frac{\lambda}{c-v} = \frac{\lambda f_s}{(c-v)f_s} = \frac{c}{(c-v)f_s} = \frac{1}{(1-\beta)f_s},$$

where f_s is the frequency of the source. The observed time, t_o, is related to the rest time, t, by Equation (1.3) as

$$t_o = \frac{t}{\gamma}.$$

The observed frequency, f_o, is therefore

$$\begin{aligned} f_o &= \frac{1}{t_o} \\ &= \frac{\gamma}{t} \\ &= \frac{1-\beta}{\sqrt{1-\beta^2}} f_s \\ &= \frac{1-\beta}{\sqrt{(1-\beta)(1+\beta)}} f_s \\ &= \sqrt{\frac{1-\beta}{1+\beta}} f_s. \end{aligned}$$

This is known as the **Doppler effect**, with the coefficient relating the frequencies being known as the **Doppler factor**.

1.8 Causality

The deeper theory of relativity is that the principle of **causality**; that is, every effect must have a preceeding cause. For such a causal universe there must be some maximum speed of information transfer. It is this that Einstein attempted to communicate in his second postulate of relativity; it is only by further reasoning and experimentation that the universal speed limit was found to be that of light in a vacuum.

1.9 Energy and Momentum

Conservation laws are paramount in every area in physics, not least of which are conservations of energy and momentum. These two quantities can be written as components of a single four-vector, known as the **four-momentum**, p. As in classical three-dimensional mechanics, the momentum in spacetime is given by the mass multiplied by the four-velocity, i.e. $p = mv$ where p and v are four-vectors. The four-momentum is therefore given by

$$\begin{aligned} p &= mv \\ &= m\gamma(c, v_x, v_y, v_z) \\ &= \gamma(mc, mv_x, mv_y, mv_z). \end{aligned}$$

At this point the time component is identified by the mass-energy equivalence $E = mc^2$ - which I assume the reader will cover at a later date and so I will not cover here - however this is only true in the rest frame of the observer. More generally the relation is $E = \gamma mc^2$, hence the four-momentum is

$$p = \left(\frac{E}{c}, \gamma m\mathbf{v} \right).$$

Chapter 2

Quantum Mechanics

"Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

— Lewis Carroll, *Alice in Wonderland*

The term "quantum mechanics" is thrown around a lot in pseudoscience to attempt to trick the public into buying products, but what does it *mean*? A "quanta" is a small amount of something, whereas "quantum" is multiple quanta, hence the term "quantum mechanics" simply means the description of motions of large numbers of small things.

2.1 History

The history of quantum physics spans back hundreds of years, to the time of Isaac Newton, to debates as to whether light is a particle ('corpuscle') or a wave. However, it is only within the past 150 years that the idea that our physical reality could, in fact, be the manifestation of waves interacting with each other and that many observable phenomena are discretised, only able to take on specific values, instead of the previously assumed continuum.

In 1877, the Austrian physicist and philosopher Ludwig Boltzmann suggested that energy levels of a system could be discretely valued, following observations of discrete energy levels within molecules. In 1900, German physicist Max Planck introduced the idea that energy is quantised in order to derive his law of blackbody radiation. Albert Einstein explained the photoelectric effect in 1905 by postulating that light, and more generally electromagnetic radiation, is divided into a "finite number of 'energy quanta' that are localised in points in space" - these energy quanta later became known as 'photons'. Niels Bohr soon explained the spectral lines of hydrogen using quantisation in 1913, and in 1923 French physicist Louis de Broglie put forward his theory of matter waves by stating that particles can exhibit characteristics of waves and vice versa. The de Broglie hypothesis was formulated in 1924 and experimentally verified in 1927 by the discovery that electrons diffract in the same manner as light. Max Born, Werner Heisenberg and Pascual Jordan developed matrix mechanics in 1925, meanwhile Erwin Schrödinger developed wave mechanics and his non-relativistic equation. In 1927, Werner Heisenberg articulated his widely-known uncertainty principle, which states that the two quantities of quantum particles cannot be known exactly simultaneously, but rather to within a limit. In the same year, English physicist Paul Dirac began unifying quantum mechanics and special relativity by combining relativity with the Schrödinger equation for the electron, producing the Dirac equation, which predicted the existence of the positron a year before it was discovered experimentally.

From 1927 onward, research in the area moved onto describing large numbers-of particles instead of individual particles/atoms, producing numerous quantum field theories. Early works were pioneered by Dirac, Pauli, Waisskopf and Jordan. During the 1940's a formulation of quantum electrodynamics was researched by Richard Feynman, Freeman Dyson, Julian Schwinger and Sin'ichirō Tomonaga and in the 1960's a quantum description of the strong interaction (quantum chromodynamics) was beginning. In 1975 this was completed by Politzer, Gross and Wilczek and work went on to unify the electromagnetic and weak forces into a quantum description of the electroweak interaction.

2.2 Units

Until now, you have probably taken SI units (Kilograms, Newtons, Joules, Watts etc.) for granted, applying them to all sorts of scenarios. Quantum phenomena are so tiny that different unit systems measuring a smaller scale become more useful, although that is not to say that quantum-scale calculations cannot be performed in SI - just that there are units better suited to the job. Nuclear physicists typically use "atomic units", accelerator physicists often implement "centimetre-gramme-second" (CGS) units, whereas particle and quantum physicists utilise "natural units" - it is these that will be used most of particle, nuclear and quantum courses.

One may ask how these units differ and what the pros and cons of it are. In short, instead of needing to factor in constant numeric quantities that arise in many equations, quantum phenomena absorbed these constants into the units, which are measured in "hbar", $\hbar$ (J s), "electronvolt", eV, and the speed of light, c.

2.3 Matter Waves, Wavefunctions and Quantum States

In order to enter the world of quantum mechanics, one must replace the Newtonian idea of matter being solid, indivisible spheres with matter instead being comprised of so-called 'matter waves'. Each of these waves can be described by its **wavefunction**.

Definition 2.1: Wavefunction

In quantum physics, a wavefunction mathematically describes a state of a system, usually a particle. The wavefunction is an extremely powerful concept as it contains information about all properties of the particle, e.g. position, momentum, energy, spin. It is usually denoted Ψ, however numerous other (usually Greek) letters are used, e.g. Φ.

The form of the wavefunction is usually complex, so cannot directly identify observable phenomena. Instead it is the distribution of the wave's **probabil-**

ity amplitudes, which can be used to gain information about the system.

Definition 2.2: Probability Amplitude

The probability amplitude, A, may be thought of as a vector or complex number corresponding to an observable phenomenon. It is not a real scalar, thus it does not indicate anything measurable by itself. However, as with vectors and complex numbers, the squared magnitude (modulus) results in a real scalar, which represents the probability of a quantum mechanical event occurring/observable being measured, i.e. $P = |A|^2 = A^* A$.

If we consider two mutually exclusive events, i.e. both cannot occur at the same time, e.g. heads and tails from a coin flip, A and B, with respective probabilities P_A and P_B, then the probability of at least one event happening - denoted $A \cup B$ - is $P_{A\cup B} = P_A + P_B$. In quantum mechanics we instead consider observable phenomena, e.g. measured values of energy, A and B, with wavefunctions Ψ_A and Ψ_B, respectively. The total amplitude, $\Psi_{A\cup B}$, is thus given by $\Psi_{A\cup B} = \Psi_A + \Psi_B$. The probability of observing either phenomenon A or B is then given by

$$\begin{aligned} P_{A\cup B} = |\Psi_{A\cup B}|^2 &= \Psi_A^* \Psi_A + \Psi_B^* \Psi_B + (\Psi_A^* \Psi_B + \Psi_B^* \Psi_A) \\ &= P_A + P_B + (\Psi_A^* \Psi_B + \Psi_B^* \Psi_A). \end{aligned}$$

This extra term indicates interference between wavefunctions, yielding different physical behaviours of the system from those predicted by classical probabilities.

A quantum system that can be described by a wavefunction, which is characterised by a set of so-called quantum numbers, is known as a **quantum state**.

Definition 2.3: Quantum State

A quantum state mathematically encodes an isolated quantum system, encapsulating the probability distributions of the outcomes of all possible measurements.

Nothing except the quantum state exists independently of observation. However, knowledge of the state as well as how it evolves with time predicts all behaviours of the system. A number of interesting pheneomena arise from the idea of states, including:

1. **Quantum "leaps"**
   Electrons can instantaneously "jump" between quantum states, from one energy level to another. For bound electrons this is known as excitation/deexcitation.

2. **Quantum superposition**
   One particle can exist in two (or more) quantum states simulatenously. The idea of Schrödinger's cat being simultaneously alive and dead is an analogy of this phenomenon.

3. **Quantum entanglement**
   Pairs of particles become inseparably intertwined such that each particle cannot be described independently of the other. This leads to unusual phenomena, for instance, a measurement of a property of one particle results in the other immediately "knowing" what measurement has been performed across arbitrarily large distances. Einstein referred to this as "spooky action at a distance".

2.4 The de Broglie Relation

The idea of matter waves is called the de Broglie hypothesis, after Louis de Broglie, who proposed a calculable wavelength, λ, associated with matter with momentum, p;

$$\lambda = \frac{h}{p}.$$

Theorem 2.1: de Broglie Wavelength

By starting from Einstein's mass-energy equivalence, namely $E = mc^2$, and Planck's theory, which states that every quantum of a wave has a discrete amount of discrete energy, $E = hf$, one can equate the energies to obtain

$$mc^2 = hf.$$

However, particles with mass cannot travel at the speed of light, and so to generalise $c \to v$;

$$mv^2 = hf.$$

Using the relation between frequency, velocity and wavelength for waves, $f = \frac{v}{\lambda}$, gives

$$mv = \frac{h}{\lambda}.$$

A final substitution of the classical momentum, $p = mv$, and rearranging gives the de Broglie wavelength as

$$\lambda = \frac{h}{p}.$$

2.5 Heisenberg's Uncertainty Principle

In short, the uncertainty principle states that the more precisely the position or momentum of a particle is known, the less precisely the other is known. It is formally written as an inequality relating the standard deviations of position, σ_x, and momentum, σ_p, such that $\sigma_x \sigma_p \geq \frac{\hbar}{2}$, where $\hbar$ is the reduced Planck's constant $\left(\hbar = \frac{h}{2\pi}\right)$.

Theorem 2.2: Heisenberg's Uncertainty Principle

Consider light of wavelength λ passing through a slit of width a which does not change the magnitude of the photon's momentum, but may change its direction. The momentum of the photon is given by the de

Broglie relation as

$$p = \frac{h}{\lambda}.$$

The momentum in the direction of the slit, say p_y, is given by $p \sin(\theta)$, where θ is the angle of diffraction of the photon. There is a distribution of maxima and minima with diffraction angle, as shown in Figure 2.1. The first minima occur at $\sin(\theta) = \pm\frac{\lambda}{a}$ and, since for small θ, $\sin(\theta) \approx \theta$, the first minima occur at $\theta \approx \pm\frac{\lambda}{a}$. Therefore $\frac{\lambda}{a}$ can be taken as the uncertainty in $\sin(\theta)$, hence the uncertainty in momentum, Δp_y, is given by

$$\begin{aligned}\Delta p_y &= p \cdot \Delta \sin(\theta) \\ &= p\left(\frac{\lambda}{a}\right).\end{aligned}$$

Substituting the de Broglie relation for momentum gives

$$\begin{aligned}\Delta p_y &= \left(\frac{h}{\lambda}\right)\left(\frac{\lambda}{a}\right) \\ &= \frac{h}{a}.\end{aligned}$$

The uncertainty in the y-position of the photon can be taken as the width of the slit, a, such that

$$\Delta y \Delta p_y = h \geq \frac{\hbar}{2}.$$

Note: This is not a thorough derivation and should not be taken as such. A more rigorous derivation of the uncertainty principle will be presented later.

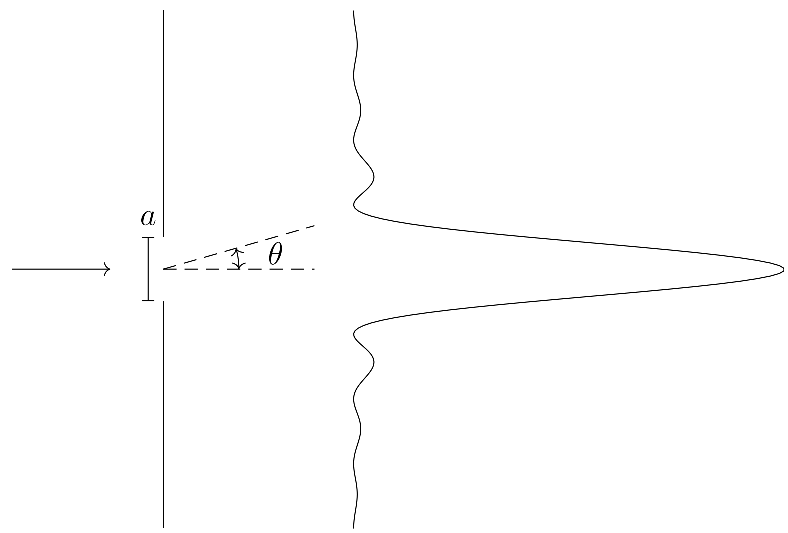

Figure 2.1: Photon angular distribution following single slit diffraction.

Example 2.1: Moving Car

Consider two people, Albert and Beth. Albert drives his car down the road while a Beth attempts to simultaneously measure his speed and location. Initially, in order to measure the speed of the car, Beth measures the distance travelled, x, in a period of time, t, and takes the ratio to get $v = \frac{x}{t}$. However, Beth didn't take into account any acceleration or deceleration - she only measured the average speed that Albert drove.

Instead, Beth decides to measure a much shorter period of time and gets a better idea of how fast Albert is driving at any point in time. However, in order to measure Albert's speed Beth requires him to move, and since he is not necessarily going a constant speed she cannot simultaneously know his location and velocity. This is the uncertainty principle on a macroscopic scale.

2.6 Applications of Quantum Mechanics

One may justly question the usefulness of quantum mechanics, given it undoes much of what is taken for granted and appears to have little correlation to the everyday macroscopic world.

Particle and nuclear physics' make extensive uses of various quantum mechanical behaviours in understanding their respective fields. Discrete electron energy levels within an atom/molecule result in spectral line spectra, allowing spectrometry across astronomical distances. The uncertainty principle can be used to predict lifetimes of radioactive substances as well as particles that are too short-lived to be directly observed. Furthermore, many-particle quantum systems exhibit group behaviours that explain observable phenomena on our scale, such as states of matter, magnetism etc.

2.7 Observables and Operators

Observables are the resultant phenomena following **operators** acting upon wavefunctions.

Definition 2.4: Observable

An observable is a quantity of a system, e.g. momentum, energy, spin.

Definition 2.5: Operator

In a broad sense of the word, an operation is the process by which a quantity of a system is brought into existence. Quantum mechanically, an operator acts upon a wavefunction to produce an observable. These are denoted with a 'hat', e.g. an operator A would be written $\hat{A}$.

The process by which operators act upon wavefunctions to produce observables remains an open question, known as the 'measurement problem'. Common interpretations include de Broglie-Bohm Theory and the Everett (Many-Worlds/Multiverse) Interpretation. The main consensus is the Copenhagen Interpretation, in which quantum systems do not have any definite properties prior to measurement, but rather the wavefunction contains a complete set of probabilities of observable values, and it is this approach that is considered for most courses on quantum mechanics.

Operators act to the right and produce an eigenvalue (observable) of the wavefunction, such that an operator $\hat{A}$ acts upon a wavefunction Ψ to produce an observable quantity a;

$$\hat{A}\Psi = a\Psi.$$

During a macroscopic experiment, the values measured are actually the average (expectation) value of the quantum bodies from which the macroscopic system is comprised. From statistics, the expectation value of a variable is given by the integral over all space of variable weighted by the probability distribution function, P. For example, in quantum mechanics $P = |\Psi|^2$ and so the measured position of a macroscopic system which is comprised of identical quantum bodies with wavefunctions Ψ is given by

$$\begin{aligned} \langle x \rangle &= \int_{-\infty}^{+\infty} Px \; dx' \\ &= \int_{-\infty}^{+\infty} \Psi^* x \Psi \; dx'. \end{aligned}$$

2.8 The Time-Independent Schrödinger Equation (TISE)

The widely-known but poorly understood Schrödinger equation is a hallmark of quantum mechanics that is usually poorly taught and rarely derived. In brief, the time-independent Schrödinger Equation (TISE) states that the Hamiltonian of a system, H, is equal to the system's total energy, E;

$$H = E.$$

However, the Hamiltonian is taken to be an operator which acts upon the system's wavefunction Ψ to produce an energy value, hence the time-independent Schrödinger Equation is

$$\hat{H}\Psi = E\Psi.$$

This will be derived and considered in great depth in later courses, but initially accepting this will simplify the transition to harder problems.

2.9 The Momentum Operator

The momentum operator, $\hat{p}$, is an operator which, when acted upon a wavefunction, results in an observation of the momentum of the system. As the derivation for the momentum operator requires further knowledge of the mathematics behind quantum mechanics, I would suggest (just for now) accepting that the momentum operator is given by

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}.$$