Contemporary Physics

Arnav Patri

April 26, 2023

Contents

Chapter 1

The Failures of Classical Physics

Under classical physics,

$$\vec{v} = \frac{\Delta \vec{r}}{t} \qquad K = \frac{1}{2}mv^2 \qquad \vec{p} = m\vec{v} \qquad \vec{L} = \vec{r} \times \vec{p} = I\vec{\omega}$$

where $\vec{v}$ is velocity, $\vec{r}$ is position, m is mass, $\vec{p}$ is momentum, $\vec{L}$ is angular momentum, and I is rotational inertia.

Working on very small scales, the unit of charge used is often the magnitude of that of an electron (or proton)

$$1\,\text{e} \approx 1.602 \times 10^{-19}\,\text{C}$$

and that of energy is the **electron-volt (eV)**, which is the energy of an electron subjected to a $1\,\text{V}$ potential difference:

$$1\,\text{eV} \approx 1.602 \times 10^{-19}\,\text{J}$$

An **atomic mass unit (amu)** is defined such that

$$12\,\text{amu} = \text{carbon atom}$$

The speed of light c and Planck's constant h are

$$c \approx 3 \times 10^8\,\frac{\text{m}}{\text{s}} \qquad h \approx 6.626 \times 10^{-34}\,\text{J}\,\text{s} \approx 4.14 \times 10^{-15}\,\text{eV}\,\text{s}$$

The product of these two constants is

$$hc \approx 1240\,\text{eV}\,\text{nm}$$

Failures of space and time require relativity while those of particle theories require quantum mechanics.

A **theory** is an organized body of facts. It provides a model of explanation.

A **hypothesis** is a prediction as to how something works. (With data to back up a hypothesis, it can become a theory.)

A **law** is a correlation between variables.

Chapter 2

The Special Theory of Relativity

2.1 Classical Relativity

A "theory of relativity" is simply a way for observers in different reference frames to compare their observations.

The mathematical basis for converting between reference frames is called a *transformation.*

Consider two observers O and O' observing the same event, O' moving relative O with constant velocity $\vec{u}$. In their own reference frames O and O' are both at rest, but they move relative to one another with constant velocity $\vec{u}$. According to O, the space and time coordinates are x, y, z, and t, while according to O', those of the *same event* are x', y', z', and t'. If the relative velocity is only in the common xx' direction, $\vec{u}$ can represent the velocity of O' as measured by O.

It is assumed that each observer is able to test and verify Newton's laws in their reference frames. A reference frame that follows Newton's first law (the law of inertia) is said to be an *inertial frame.* In order for a frame to be inertial, it must not be accelerating.

The classical *Galilean* transformation relating the coordinates assumes as a postulate of classical physics that $t = t'$; that is, time is the same for all observers. It is also assumed for simplicity that the coordinate systems are chosen such that their origins overlap at $t = 0$. Consider an object in O' at coordinates (x', y', z'). According to O, the y and z coordinates are the same as those observed in O'. Along the x direction, though, O would observe the object at $x = x' + ut$. This yields the *Galilean coordinate transformation*

$$x' = x - ut \qquad y' = y \qquad z' = z \qquad t' = t \qquad \text{(Galilean coordinate transformation)}$$

The velocities as observed by O can be found simply by differentiating the prior results:

$$v'_x = v_x - u \qquad v'_y = v_y \qquad v'_z = v_z \qquad \text{(Galilean velocity transformation)}$$

Again differentiating yields the relationships between the accelerations:

$$a'_x = a_x \qquad\qquad a'_y = a_y \qquad\qquad a'_z = a_z$$

This shows that Newton's laws hold for both observes so long as u is constant, as the observers measure identical accelerations, meaning that the results of applying $\vec{F} = m\vec{a}$ are also identical.

2.3 Einstein's Postulates

1. **Principle of Relativity** The laws of physics are identical for all inertial reference frames.

2. The speed of light is constant for all inertial reference frames.

2.4   Consequences of Einstein's Postulates

As a consequence of these two postulates, the motion of the observer determines their measurements.

The Relativity of Time

Consider a beam of light traveling up to a mirror and back down (in the y direction) in a reference frame moving at speed u in the x direction. An observer. An observer O at rest and an observer O' moving in the same reference frame as the mirror.

From the perspective of O', the time interval between the light leaving and returning from its source is

$$\Delta t_0 = \frac{2L_0}{c}$$

where L_0 is the distance between the origin of the light and the mirror.

From the perspective of O, the light beam travels a distance $2L$, where

$$L = \sqrt{L_0^2 + \left(\frac{u\Delta t}{2}\right)^2}$$

as derived by Pythagorean theorem. As O observes the beam of light to travel at c, the time interval measured is

$$\Delta t = \frac{2L}{c} = \frac{2\sqrt{L_0^2 + (u\Delta t/2)^2}}{c}$$

Substituting for L_0 and solving for Δt yields

$$\Delta t = \frac{\Delta t_0}{\sqrt{1 - u^2/c^2}} \qquad \text{(time dilation)}$$

As the $u \leq c$, the denominator of this expression is at most 1, meaning that $\Delta t \leq \Delta t_0$; that is, O measures a longer time interval than O'. For observer O, the beginning and ending of the time interval occur at the same location, so the interval Δt_0 observed by them is known as *proper time*.

The Relativity of Length

Suppose the beam of light instead travels in the same direction as that of the motion of O'. According to O, the distance traveled is L, while from the perspective of O' (from which the clock is at rest), it is L_0.

The light takes time interval Δt_1 to reach the mirror, in which it travels distance $c\Delta t_1$, which is equal to the length L plus the distance $u\Delta t_1$ traveled by by O':

$$c\Delta t_1 = L + u\Delta t_1$$

The light then takes time interval Δt_2 to return to its source, in which it travels distance $c\Delta t_2$, equal to the length L minus the distance $u\Delta t_2$:

$$c\Delta t_2 = L - u\Delta t_2$$

Adding these times yields Δt:

$$\Delta t = \Delta t_1 + \Delta t_2 = \frac{L}{c-u} + \frac{L}{c+u} = \frac{2L}{c}\frac{1}{1-u^2/c^2}$$

Applying the formula for time dilation,

$$\Delta t = \frac{\Delta t_0}{\sqrt{1-u^2/c^2}} = \frac{2L_0}{c}\frac{1}{\sqrt{1-u^2/c^2}}$$

Setting the above two equations equal to each other yields

$$L = L_0\sqrt{1 - \frac{u^2}{c^2}} \qquad\qquad \text{(length contraction)}$$

Observer O', who is at rest relative to the object, measures the *rest length* L_0 (the *proper length*) while all observers relative to whom O' is in motion measure a shorter length *only in the direction of motion*. A moving body appears to be shorter than one at rest.

It should be noted that length contraction occurs only in the direction of movement, meaning that any dimensions perpendicular to that of movement are unaffected.

Relativistic Velocity Addition

A source emits particles traveling at speed v' according to observer O', who is at rest with respect with respect to the source. A bulb is triggered to flash when the particle reaches it. This flash makes a return trip to the source. The time interval Δt_0 measured by O' is comprised of the time it takes for the particle to travel distance L_0 at speed v' and the time it takes the flash to travel the same distance at c:

$$\Delta t_0 = \frac{L_0}{v'} + \frac{L_0}{c}$$

O' is moving at speed u relative to observer O. For O, the particle traveling at speed v according to O reaches the bulb after Δt_1, traveling a distance of $v\Delta t_1$ equal to the (contracted) length L plus the distance $u\Delta t_1$ moved by O':

$$v\Delta t_1 = L + u\Delta t_1$$

In the interval Δt_2, the light beam traverses distance $c\Delta t_2$ equal to length L minus the distance $u\Delta t_2$ traveled by O':

$$c\Delta t_2 = L - u\Delta t_2$$

Solving the above two equation for Δt_1 and Δt_2 and adding them yields the total interval Δt between the emission of the particle and the arrival of the beam at the detector according to O. Applying the time dilation formula to relate this result to Δt_0 and using the length contraction formula to relate L to L_0 yields the result

$$v = \frac{v'+u}{1+v'/uc^2} \qquad\qquad \text{(relativistic velocity addition)}$$

Note that when $v' = c$ (a beam of light being observed),

$$v = \frac{c+u}{1+cu/c^2} = \frac{c+u}{1+u/c} = \frac{c+u}{\dfrac{c+u}{c}} = c$$

regardless of the value of u. All observers measure the same speed of light c regardless of their reference frame.

The Relativistic Doppler Effect

The classical Doppler effect states that an observer O moving relative to the source of waves detects a frequency f' different from the true frequency f being emitted by the source S:

$$f' = f\frac{v \pm v_O}{v \mp v_S} \qquad \text{(classical Doppler effect)}$$

where v is the speed of the waves through some medium, v_S is the speed of the source *relative to the medium*, and O is that of the observer *relative to the medium*. The upper signs are chosen when S and O are moving towards each other while the lower ones are chosen when they are moving away from each other.

Consider a source of electromagnetic waves at rest in the reference frame of observer O. An observer O' moving relative to the source at speed u will observe a different frequency. Suppose O observes N waves emitted at frequency f. It takes an interval $\Delta t_0 = N/f$ for these N waves to be emitted from the point of view of O. This is the proper time interval in the frame of reference of O. The corresponding interval to O' is $\Delta t'$, during which O moves a distance $u\Delta t'$. The wavelength λ' according to O' is the total length occupied by the waves divided by the number of waves:

$$\lambda' = \frac{c\Delta t' + u\Delta t'}{N} = \frac{c\Delta t' + u\Delta t'}{f\Delta t_0}$$

The frequency from the reference frame of O' is $f' = c/\lambda'$, so

$$f' = \frac{c}{\lambda'} = f\frac{\Delta t_0}{\Delta t'}\frac{1}{1 + u/c}$$

Applying the time dilation formula,

$$f' = f\frac{\sqrt{1 - u^2/c^2}}{1 + u/c} = f\sqrt{\frac{1 - u/c}{1 + u/c}} \qquad \text{(relativistic Doppler shift)}$$

This is the formula for the *relativistic Doppler shift* where the waves are observed parallel to $\vec{u}$. It should be noted that unlike the classical formula, this formula does *not* distinguish between the source and observer motion; it is only dependent on the relative speed u between the source and observer.

Restated,

$$\lambda' = \lambda\sqrt{\frac{1 + u/c}{1 - u/c}} \qquad\qquad \frac{u}{c} = \frac{(\lambda'/\lambda)^2 - 1}{(\lambda'/\lambda)^2 + 1}$$

2.5   The Lorentz Transformation

While the Galilean transformation holds for the "common-sense" experience observed at low speeds, it does not agree with experiments performed at high speeds. A new set of transformations is therefore required that is capable of reproducing relativistic effects such as time dilation, length contraction, velocity addition, and the Doppler shift.

A transformation that enables observers O and O' moving relative to each other to compare their measurements in space and time coordinates of the same event. Teh transformation equations

relate the measurements x, y, z, t in O to x', y', z, and t' in O'. This transformation must be linear (depending only on the first powers of each coordinate, following the homogeneity of space and time), consistent with Einstein's postulates, and it must reduce to the Galilean transformation when the relative speed between the observers is small. It is again assumed that the relative speed is only in the positive xx' direction.

The new transformation is called the *Lorentz transformation*:

$$x' = \frac{x - ut}{1 - u^2/c^2} \qquad y' = y \qquad z' = z \qquad t' = \frac{t - (u/c^2)x}{\sqrt{1 - u^2/c^2}} \qquad \text{(Lorentz transformation)}$$

This is sometimes rewritten as

$$x' = \gamma(x - \beta ct) \qquad y' = y \qquad z' = z \qquad ct' = \gamma(ct - \beta x)$$

or

$$x' = \gamma(x - vt) \qquad y' = y \qquad z' = z \qquad t' = \gamma\left(t - x\frac{v}{c^2}\right)$$

or in matrices

$$\begin{bmatrix} x' \\ y' \\ z' \\ ct' \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix} \qquad \begin{bmatrix} x' \\ ct' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix} \qquad \begin{bmatrix} x \\ ct \end{bmatrix} = \begin{bmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} x' \\ ct'' \end{bmatrix}$$

where

$$\beta = \frac{u}{c} \qquad \text{and} \qquad \gamma = \frac{1}{\sqrt{1 - u^2/c^2}}$$

It is often useful to write these equation in terms of intervals by replacing each coordinate by their corresponding interval.

The equations for y' and z' are identical to their Galilean counterparts while the first and last equations reduce to theirs when $u \ll c$.

The Lorentz transformation can then be used to derive some predictions of special relativity.

Length Contraction

A rod of length L_0 is at rest as observed by O'. The rod extends from x_1' to x_2'; that is, $L_0 = x_2' - x_1'$. Observer O, relative to whom the rod is in motion, measures the ends of the rod at x_1 and x_2. For the length of the rod to be measured by O, x_1 and x_2 must be *simultaneously* determined.

Suppose O' flashes a bulb at x_1' and t_1' and then flashes a bulb at x_2' and t_2'. The Lorentz transformation gives

$$x_1' = \frac{x_1 - ut_1}{\sqrt{1 - u^2/c^2}} \qquad\qquad x_2' = \frac{x_2 - ut_2}{\sqrt{1 - u^2/c^2}}$$

Subtracting the first equation from the second yields

$$x_2' - x_1' = \frac{x_2 - x_1}{\sqrt{1 - u^2/c^2}} - \frac{u(t_2 - t_2)}{\sqrt{1 - u^2/c^2}}$$

9

The bulbs must be set off such that the flashes appear to be simultaneous to O. This enables a simultaneous determination to be made of the rod's endpoints. if the flashes are observed simultaneously, $t_1 = t_2$, so

$$x_2' - x_1' = \frac{x_2 - x_1}{\sqrt{1 - u^2/c^2}}$$

Substituting L_0 for $x_2' - x_1'$ and L for $x_2 - x_1$ yields

$$L = L_0\sqrt{1 - u^2/c^2}$$

Velocity Transformation

Suppose O observes a particle traveling with velocity $\vec{v}$. The relationship between $\vec{v}$ and the velocity $\vec{v}'$ observed by O' is given by the *Lorentz velocity transformation*

$$v_x' = \frac{v_x - u}{1 - v_x u/c^2} \quad v_y' = \frac{v_y\sqrt{1 - u^2/c^2}}{1 - v_x u/c^2} \quad v_z' = \frac{v_z\sqrt{1 - u^2/c^2}}{1 - v_x u/c^2} \quad \text{(Lorentz velocity transformation)}$$

Solving the above equations for v_x, it can be shown that this is identical to the equation for relativistic velocity addition. It should be noted that for $u \ll c$, the Lorentz velocity transformation reduces to its Galilean counterpart. Note also that despite $y' = y$, $v_y' \neq v_y$. This is due to the Lorentz transformation's handling of t.

These transformation can be derived from the coordinate transformation. Differentiating the coordinate transformation for x' yields

$$dx' = \frac{dx - u\,dt}{1 - u^2/c^2}$$

Differentiating the time coordinate transformation yields

$$dt' = \frac{dt - (u/c^2)\,dx}{1 - u^2/c^2}$$

v_x' is therefore

$$v_x' = \frac{dx'}{dt'} = \frac{\dfrac{dx - u\,dt}{1 - u^2/c^2}}{\dfrac{dt - (u/c^2)\,dx}{1 - u^2/c^2}} = \frac{dx - u\,dt}{dt - (u/c^2)\,dx} = \frac{dx/dt - u}{1 - (u/c^2)\,dx/dt} = \frac{v_x - u}{1 - v_x u/c^2}$$

Differentiating $y' = y$ yields $dy' = dy$, so

$$v_y' = \frac{dy'}{dt'} = \frac{dy}{\dfrac{dt - (u/c^2)\,dx}{\sqrt{1 - u^2/c^2}}} = \frac{dy\sqrt{1 - u^2/c^2}}{dt - (u/c^2)\,dx} = \frac{\sqrt{1 - u^2/c^2}\,dy/dt}{1 - (u/c^2)\,dx/dt} = \frac{v_y\sqrt{1 - u^2/c^2}}{1 - uv_x/c^2}$$

Differentiating for $z' = z$ yields $dz' = dz$, so

$$v_z' = \frac{dz'}{dt'} = \frac{dz}{\dfrac{dt - (u/c^2)\,dx}{\sqrt{1 - u^2/c^2}}} = \frac{dz\sqrt{1 - u^2/c^2}}{dt - (u/c^2)\,dx} = \frac{\sqrt{1 - u^2/c^2}\,dz/dt}{1 - (u/c^2)\,dx/dt} = \frac{v_z\sqrt{1 - u^2/c^2}}{1 - uv_x/c^2}$$

2.6 The Twin Paradox

Suppose there are a pair of twins on earth. Twin A remains on Earth while twin B sets off on a rocket ship. A believes that as a consequence of time dilation, B should be younger than them by the time they return. For two observers in relative motion, though, *each* believes the *other's* clock to be slow. From the point of view of B, A make a round-trip away from them and back again. Under such circumstances, it should be A's clock that is running slow, so A they expect A to be younger when they return.

The resolution of this paradox comes from considering the asymmetric roles of the twins. Special relativity applies only to inertial frames. In order to leave and reenter Earth, B must accelerate and correspondingly decelerate. Although this may be over a brief time interval, B's return journey occurs in a different inertial frame to their departure. This change in B's inertial frame is what results in the asymmetry in the twins' ages. Only B is required to change inertial frames to return, so *all observers will agree* that B is the one that is "really" in motion, so it is therefore their clock that is "really" running slow, meaning that B will indeed be the younger twin upon their return.

Spacetime Diagrams

A *spacetime* diagram plots time on the y-axis and space on the x-axis. The slope of a straight line is velocity as a proportion of c.

2.7 Relativistic Dynamics

Dynamical quantities such as momentum and kinetic energy are dependent on length, time, and velocity. As such, they are also affected by Einstein's postulates.

Consider a collision of two particles that is elastic from the reference frame of O'. Particle 1 of mass $m_1 = 2m$ is initially at rest while particle 2 of mass $m_2 = m$ is moving in the negative x direction with initial velocity $v'_{2,i} = -0.75c$. The final velocities of particles 1 and 2 are respectively $v'_{1,f} = -0.5c$ and $v'_{2,f} = 0.25c$. Analyzing this collision classically,

$$p'_i = m_1 v'_{1,i} + m_2 v'_{2,i} = -0.75mc \qquad\qquad p'_i = m_1 v'_{2,f} + m_2 v'_{2,f} = -0.75mc = p'_i$$

As the initial and final momenta are equal according to O', energy is conserved.

Suppose the reference frame of O' is moving in the positive x direction at a speed of $u = 0.55c$ relative to O. The velocity transformation yields the velocities of $v_{1,i} = 0.55c$, $v_{2,i} = -0.34c$, $v_{1,f} = 0.069c$, and $v_{2,f} = 0.703c$. Analyzing the collision,

$$p_i = m_1 v_{1,i} + m_2 v_{2,i} = 0.76mc \qquad\qquad p_f = m_1 v_{1,f} + m_2 v_{2,f} = 0.841mc \neq p_i$$

As $p_i \neq p_f$, momentum is not conserved according to O. This means that the law of conservation of linear momentum does not satisfy Einstein's first postulate. To retain the conservation of momentum, then, a new definition of momentum is required. This definition must yield a conservation law that holds under relativity (momentum being conserved regardless of the reference frame, so long as it is inertial) and reduces to $\vec{p} = m\vec{v}$ at low speeds.

These requirements are satisfied by defining relativistic momentum to be

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - v^2/c^2}} \qquad\qquad \text{(relativistic momentum)}$$

The velocity v is the velocity of the particle as measured from a particular inertial frame; it is not the velocity of said frame.

Using this relativistic definition of momentum to analyze the above collision,

$$p'_i = \frac{m_1 v'_{1,i}}{\sqrt{1 - v'^2_{1,i}/c^2}} + \frac{m v'_{2,i}}{\sqrt{1 - v'^2_{2,i}/c^2}} \approx 1.134mc \quad p'_f = \frac{m_2 v'_{1,f}}{\sqrt{1 - v'^2_{1,f}/c^2}} + \frac{m_2 v'_{2,f}}{\sqrt{1 - v'^2_{2,f}/c^2}} \approx 1.134mc = p'_i$$

As $p'_i = p'_f$, O' is able to conclude that momentum is conserved. To O,

$$p_i = \frac{m_1 v_{1,i}}{\sqrt{1 - v^2_{1,i}/c^2}} + \frac{m v_{2,i}}{\sqrt{1 - v^2_{2,i}/c^2}} \approx 0.956mc \quad p_f = \frac{m_1 v_{1,f}}{\sqrt{1 - v^2_{1,f}/c^2}} + \frac{m_2 v_{2,f}}{\sqrt{1 - v^2_{2,f}/c^2}} \approx 0.946mc = p_i$$

so O' can also conclude that momentum is conserved.

Relativistic Kinetic Energy

Much like classical momentum, classical kinetic energy also fails under relativity. According to O', the aforementioned collision has kinetic energies

$$K'_i = \frac{1}{2}m_1 v'^2_{1,i} + \frac{1}{2}m_2 v'^2_{2,i} \approx 0.281mc^2 \qquad K'_f = \frac{1}{2}m_1 v'^2_{1,f} + \frac{1}{2}m_2 v'^2_{2,f} \approx 0.281mc^2 = K'_i$$

but according to O is has

$$K_i = \frac{1}{2}m_1 v^2_{1,i} + \frac{1}{2}m_2 v^2_{2,i} \approx 0.36mc^2 \qquad K_f = \frac{1}{2}m_1 v^2_{1,f} + \frac{1}{2}m_2 v^2_{2,f} \approx 0.252mc^2 \neq K_i$$

so O concludes that energy is not conserved.

The relativistic definition of kinetic energy is

$$K = \frac{mc^2}{\sqrt{1 - v^2/c^2}} - mc^2 \qquad \text{(relativistic kinetic energy)}$$

There is no limit to the amount of kinetic energy that can be given to a particle. Under the classical definition of kinetic energy, this means that speed must also lack a limit, which violates Einstein's second postulate.

Relativistic Total Energy and Rest Energy

Relativistic kinetic energy can also be expressed as

$$K = E - E_0 \qquad \text{(relativistic kinetic energy)}$$

where the *relativistic total energy* E is defined as

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \qquad \text{(relativistic total energy)}$$

and the *rest energy* E_0 is defined is

$$E_0 = mc^2 \qquad \text{(rest energy)}$$

That of en electron is
$$E_{\mathrm{e}} \approx 0.511 \, \mathrm{MeV}$$

This suggests that mass can be expressed in units of MeV/c^2. The m in this equation is sometimes called the *rest mass* m_0, which is distinguished from the "relativistic mass" m', defined as
$$m' = \frac{m_0}{\sqrt{1 - v^2/c^2}} \qquad \text{(relativistic mass)}$$

The total energy can be found as the sum of the kinetic and rest energies:
$$E = E_0 + K$$

Manipulating the equations for relativistic momentum and total energy yields the relationship
$$E = \sqrt{(pc)^2 + (mc^2)^2}$$

This can be remembered by
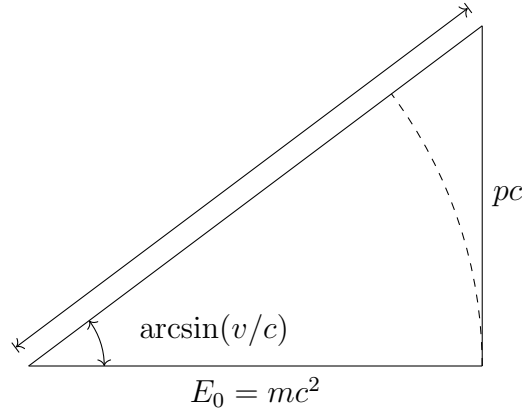
A particle traveling close to the speed of light $(v > 0.99c)$, $K \gg E_0$. IN this case,
$$E \cong pc$$

This is the *extreme relativistic approximation* and can often simplify calculations. For massless particles (such as photons),
$$E = pc$$

All massless particles travel at c; otherwise, their kinetic and total energies would be zero.

2.8 Conservation Laws in Relativistic Decays and Collisions

The law of conservation of momentum can be applied to all collisions. The only difference for high speeds is that the relativistic definition of momentum must be used. It can be stated the same:

> *In an isolated system of particles, the total linear momentum remains constant.*

Classically, the only form of energy present in an elastic collision is kinetic, so the conservation of energy is simply that of kinetic energy. In inelastic collisions or decay processes, kinetic energy is not conserved. The total energy does indeed remain constant, though; the other forms of energy are simply unaccounted for. This missing energy is usually stored in particles as atomic or nuclear energy.

Relativistically, the internal stored energy is accounted for by the rest energy. In atomic and nuclear processes, the only forms of energy generally accounted for are kinetic and rest (though radiation will be considered later). A loss of kinetic energy then implies an increase in rest energy, as the total relativistic energy does not change.

Consider a reaction in which new particles are produced. The loss of kinetic energy in the original particles is what provides the rest energy of the new particle. In a nuclear decay process, on the other hand, the initial nucleus loses some rest energy in the form of the kinetic energy given to the decay products.

The relativistic law of conservation of energy is as follows:

> *In an isolated system of particles, the relativistic total energy (kinetic energy plus rest energy) remains constant.*

When applying this law to relativistic collisions, it does not matter whether a collision is elastic or inelastic, as the inclusion of rest energy accounts for any potential loss of kinetic energy.

2.9   Extra

The **Epstein Cosmic Speedometer** is a plot of proper time (on the vertical axis) against space (on the horizontal axis). A body in motion will move through both space and time, making it an angle line. A stationary body will move only in time, making it a vertical line. The only thing able to move through space but not time is light, making it a horizontal line.

A spacetime diagram maintains the length of all vectors, plotting space on the horizontal axis and time on the vertical axis. The vertical component of a vector is the "proper" while its horizontal component is the speed through space.

Moving through space means moving slower through time; that is, moving objects effectively have slower clocks.

Graphs can be made in Lorentz/Minkowski Space. The axes are x and ct. The new axes can be found as lines with direct variation of slopes x' and ct'. Note that ct' and x' are not perpendicular. To graph with them, then, one must first go along one axis and then parallel along the other. The greater β, is, the closer together the ct' and x' axes; that is, the slower a reference frame is traveling, the less it skews spacetime. The graph can be used to plot a point in one reference frame and see its coordinates in the other.

It should be noted that

$$(ct)^2 - x^2 = (ct')^2 - x'^2 \qquad \text{(invariant)}$$

Chapter 3

The Particlelike Properties of Electromagnetic Radiation

Wave mechanics is the second theory on which modern physics is based. One of its consequences is the breakdown of the classically made distinction between particles and waves. Light, which is usually treated as a wave, can have properties associated with particles. Rather than spreading energy smoothly over a wave, it is delivered in discrete packets (*quantum*) known as *photons*.

3.1   Review of Electromagnetic Waves

An electromagnetic field is characterized by its electric field $\vec{E}$ and magnetic field $\vec{B}$. The electric field at distance r from a point charge q is

$$\vec{E} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}\hat{r} \qquad \text{(electric field)}$$

where $\hat{r}$ is a unit vector in the radial direction. The magnetic field at distance r from a long, straight wire carrying a current i along the z-axis is

$$\vec{B} = \frac{\mu_0 i}{2\pi r}\hat{\varphi} \qquad \text{(magnetic field)}$$

where $\hat{\varphi}$ is a unit vector in the azimuthal direction (in the xy-plane) in cylindrical coordinates. (Note this definition is different than its mathematical definition, where the azimuthal angle is in the yz-plane.)

If the charges are accelerated or the current is varied with time, an electromagnetic wave is produced, in which the electric and magnetic fields vary not only with $\vec{r}$ but also with t. The mathematical expression used to describe this can take several forms depending on the properties of the source wave and the medium that is being traversed. A special form is the *plane wave*, for which the wave fronts are planes. (A point source produces spherical waves, having spherical wave fronts.) A plane electromagnetic wave traveling in the positive z direction is described by

$$\vec{E} = \vec{E}_0 \sin(kz - \omega t) \qquad\qquad \vec{B} = \vec{B}_0 \sin(kz - \omega t) \qquad \text{(planar wave)}$$

where the *wave number* k is found as

$$k = \frac{2\pi}{\lambda} \qquad \text{(wave number)}$$

where λ is the wavelength and the *angular frequency* ω is found as

$$\omega = 2\pi f \qquad \text{(angular frequency)}$$

where f is the frequency. As $c = \lambda f$,

$$c = \frac{\omega}{k}$$

The wave's polarization is represented by $\vec{E}_0$, the plane of polarization being determined by the direction of $\vec{E}_0$ and the direction of propagation (in this case $+z$). The direction of $\vec{B}_0$ is then fixed by the requirement that $\vec{B}$ be perpendicular to $\vec{E}$ and the direction of travel and that $\vec{E} \times \vec{B}$ must point in the direction of travel.

The magnitude of $\vec{B}_0$ is determined by

$$B_0 = \frac{E_0}{c}$$

An electromagnetic wave transfers energy, the flux of which is determined by the *Poynting vector* $\vec{S}$:

$$\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B} \qquad \text{(Poynting vector)}$$

For a plane wave, this reduces to

$$\vec{S} = \frac{1}{\mu_0} E_0 B_0 \sin^2(kz - \omega t) \; \hat{k}$$

The units of the Poynting vector are that of power per unit area, such as W/m^2.

An electromagnetic wave with $\vec{E}$ oscillating about the x-axis and $\vec{B}$ about the y, $\vec{P}$ points in the $+z$ direction.

Consider a detector of electromagnetic radiation with area A placed perpendicular to the z axis. The power entering the receiver is

$$P = SA = \frac{1}{\mu_0} E_0 B_0 A \sin^2(kz - \omega t)$$

which can be rewritten as

$$P = \frac{1}{\mu_0 c} E_0^2 A \sin^2(kz - \omega t)$$

Note that the intensity (the average power per unit area) is proportional to E_0^2. In general, *the intensity of a wave is proportional to the square of its amplitude.*

It should also be noted that the intensity fluctuates with time with frequency

$$f = \frac{\omega}{2\pi}$$

This rapid fluctuation is typically not observed. The average power can be found from the observation time T as

$$P_{\text{avg}} = \frac{1}{T} \int_0^T P \, dt$$

The intensity I can be obtained as

$$I = \frac{P_{\text{avg}}}{A} = \frac{1}{2\mu_0 c} E_0^2$$

as the average value of $\sin^2 \theta$ is 0.5.

16

Interference and Diffraction

The most unique property of waves as physical phenomena is the *principle of superposition*, which allows two waves meeting at a point to have a combined disturbance on that point that may be greater or less than that provided each wave individually.

This property leads to *interference* and *diffraction*.

The simplest example of the former is *Young's double-slit experiment.* A monochromatic plane wave is given two narrow slits. The slits *diffract* the plane wave, resulting in the area covered by the light passing through each slit being far greater than the areas of the slits. This causes the light from each slit to overlap, resulting in an interference pattern.

A bright region in the interference pattern is from *constructive interference*, where the peaks of each beam align, resulting in their amplitudes being added. It occurs at point where the distance from one slit is exactly one wavelength greater than that from the other.

If X_1 and X_2 are distances from the point to the two slits, then in order for maximum interference to occur

$$|X_1 - X_2| = n\lambda, \quad n \in \mathbb{Z}^+ \qquad \text{(constructive interference)}$$

When the two waves cancel, there is a dark region, created by *destructive interference.* (The existence of destructive interference shows that the powers are not added, as power is always positive, but rather the electric fields.) This occurs when the distances are such that the phases of the waves differ by an integer number of half-cycles:

$$|X_1 - X_2| = \frac{2n+1}{2}, \quad n \in \mathbb{N} \qquad \text{(destructive interference)}$$

Crystal Diffraction of X Rays

3.2 The Photoelectric Effect

When a light is shone on a metal surface, electrons may be emitted. This is know as the *photoelectric effect.* The emitted electrons are called *photoelectrons.*

The rate of electron emission from the surface can be measured as an electric current by an ammeter. The maximum kinetic energy of the electrons can then be measured by applying a negative potential V_S (called the *stopping potential* to a collector for the electrons that is just enough to repel the most energetic electrons (the ammeter dropping to 0). This means that

$$K_\text{max} = -\Delta U = eV_\text{s}$$

Classically, the metal's surface is illuminated by en electromagnetic wave of intensity I. The surface absorbs the energy until that energy exceeds the binding energy of the electron, at which point the electron is released. This minimum energy is called the metal's *work function* φ.

The Classical Theory of the Photoelectric Effect

Classical wave theory makes three key predictions regarding the emitted photoelectrons:

1. *The maximum kinetic energy of the electrons should be proportional to the intensity of the radiation.* As the brightness increases, so does the energy carried by the radiation, meaning that more energy is transferred to the surface (the electric field is greater), so the kinetic energy of the electrons should also be greater.

2. *The photoelectric effect should occur for light of any frequency or wavelength.* So long as the light is intense enough to release electrons, the photoelectric effect should occur regardless of the frequency or wavelength.

3. *The first electrons should be emitted on the order of a few seconds from the radiation beginning to strike the surface.* The energy of the wave is uniformly distributed over the wave front. If the energy is absorbed directly from the wave, the amount of energy emitted to an electron is dependent on how much radiant energy is incident on the area in which the electron is confined. Assuming this area to be about the size of an atom, the time lag should be on the order of seconds.

The Quantum Theory of the Photoelectric Effect

Einstein proposed that the energy of electromagnetic radiation is not continuously distributed but rather in discrete packets called *quanta* (or *photons*). A photon's energy is associated with the wavelength and frequency as

$$E = hf = \frac{hc}{\lambda} \qquad \text{(energy of a photon)}$$

Photons are often treated as particles. Like electromagnetic waves, they travel at the speed of light. They must also obey the relativistic relationship $p = E/c$, which means that

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

In addition to energy, photons carry linear momentum.

Because a proton travels at the speed of light, its mass must be 0; otherwise, it would have infinite energy and momentum. Its rest energy is also 0.

According to Einstein, a photoelectron is released as a result of an encounter with a *single photon*, the entire energy of which is delivered instantaneously to a *single photoelectron*. If the photon energy is greater than the work function of the material, the photoelectron will be released. Otherwise, the photoelectric effect will not occur. This explanation accounts for two failures of classical wave theory: the existence of a cutoff frequency and the lack of a measurable time delay.

If the photon's energy exceeds the work function, the excess energy becomes kinetic energy for the electron:

$$K_{\max} = hf - \varphi$$

The intensity of the light is not factored into this equation. Increasing the intensity while maintaining the frequency and wavelength simply means that the number of photoelectrons released is increased; their maximum kinetic energies are unchanged.

A photon with energy equal to the work function corresponds with light of frequency f_c, called the cutoff frequency:

$$f_c = \frac{\varphi}{h} \qquad \text{(cutoff frequency)}$$

The corresponding cutoff wavelength is

$$\lambda_c = \frac{hc}{\varphi} \qquad \text{(cutoff wavelength)}$$

This is the *largest* wavelength for which the cutoff frequency can be observed.

3.3 Thermal Radiation

Thermal radiation, the radiation released by all objects due to their temperature, is not explained by classical wave theory.

Experiments yield two key characteristics regarding radiation:

1. The total intensity radiated over all wavelengths (the area under each curve) increases with temperature as

$$I = \sigma T^4 \qquad \text{(Stefan's law)}$$

   where proportionality constant σ is

$$\sigma \approx 5.67 \times 10^{-8} \, \frac{\text{W}}{\text{m}^2} \, \text{K}^4 \qquad \text{(Stefan-Boltzmann constant)}$$

   as determined experimentally.

2. The wavelength λ_{max} at which the intensity peaks decreases as the temperature increases:

$$\lambda_{\text{max}} T \approx 2.9 \times 10^{-3} \, \text{m K} \qquad \text{(Wien's displacement law)}$$

To simplify analysis of thermal radiation, a *blackbody* is considered, which absorbs all radiation incident on it, reflecting none of it. To further simplify, a special type is considered: a hole in a hollow metal box with walls that are in thermal equilibrium at temperature T. The box is filled with electromagnetic radiation that is emitted and reflected by the walls. A small hole in one of the walls allows some radiation to escape. *It is the hole, not the box itself, that is the blackbody.* Radiation from outside that is incident on the hole enters the box, having a negligible change of reemerging from the hole; hence no reflections occur from the blackbody. The radiation that emerges from the hold is a sample of the radiation inside the box, so understanding the radiation inside the box enables an understanding of the radiation leaving through the hole.

Consider the radiation inside the box. It has energy density per unit wavelength interval $u(\lambda)$; that is, the energy density of the electromagnetic radiation with wavelengths between λ and $\lambda + \mathrm{d}\lambda$ in a small volume element is $u(\lambda) \, \mathrm{d}\lambda$. At a given instant, half of the radiation within the box is moving away from the hole. The other half is moving towards it with velocity of magnitude c directed in a range of angles. Averaging over this range to evaluate the energy flowing perpendicular to the hole's surface introduces another factor of $1/2$, making the contribution of the radiation in this wavelength interval to the intensity passing through the hole

$$I(\lambda) = \frac{c}{4} u(\lambda)$$

The quantity $I(\lambda) \, \mathrm{d}\lambda$ is the radiant intensity over the interval $\mathrm{d}\lambda$. To find the total intensity emitted, one must integrate this quantity over all wavelengths:

$$I = \int_0^\infty I(\lambda) \, \mathrm{d}\lambda$$

Classical Theory of Thermal Radiation

The following predictions are made by classical theories of electromagnetism and thermodynamics regarding the relationship between I and λ.

1. *The box is filled with electromagnetic standing waves.* As the box's walls are metal, the radiation is reflected back and forth with a node of the electric field at each wall (the electric field within a conductor being 0). This same condition applies to other standing waves.

2. *The number of standing waves with wavelengths between λ and $\lambda + \mathrm{d}\lambda$ is*

$$N(\lambda)\,\mathrm{d}\lambda = \frac{8\pi V}{\lambda^4}\,\mathrm{d}\lambda$$

   where V is the volume of the box. For one-dimensional standing waves, as on a string of length L, the allowed wavelengths are

$$\lambda = \frac{2L}{n} \quad \text{for} \quad n \in \mathbb{Z}^+$$

   The number of possible standing waves with wavelengths between λ_1 and λ_2 is

$$n_2 - n_1 = 2L\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right)$$

   In the interval from λ to $\mathrm{d}\lambda$, the number of standing waves is

$$N(\lambda)\,\mathrm{d}\lambda = \left|\frac{\mathrm{d}n}{\mathrm{d}\lambda}\right|\mathrm{d}\lambda = \frac{2L}{\lambda^2}\,\mathrm{d}\lambda$$

   Extending this approach to three dimensions yields the number of standing waves between λ and $\lambda + \mathrm{d}\lambda$ as stated above.

3. *Each individual waves contributes an average energy kT to the radiation in the box.*

Quantum Theory of Thermal Radiation

3.4   The Compton Effect

The Compton effect describes the phenomenon of radiation scattering from loosely bound, nearly free electrons. Part of the energy of the radiation is given to the electron while the remainder is reradiated as electromagnetic radiation. Under classical wave theory, the scattered radiation is less energetic than the incident radiation (as some energy must go into the kinetic energy of the electron). The concept of photons, however, gives rise to a much different prediction.

The scattering process is analyzed as an interaction between a single photon and electron, which is assumed to be at rest. Initially, the photon has energy E and linear momentum p given by

$$E = hf = \frac{hc}{\lambda} \qquad \text{and} \qquad p = \frac{E}{c}$$

The electron has rest energy $m_e c^2$. After the scattering, the photon has energy and momentum

$$E' = \frac{hc}{\lambda'} \qquad \text{and} \qquad p' = \frac{E'}{c}$$

and is moving in a direction at angle θ with respect to that of the incident photon. The electron has total final energy E_e and momentum p_e and moves in a direction at angle φ with respect to the

initial photon. (To account for high-energy incident photons given energetic scattering electrons, relativistic kinematics are used for the electron.) Applying the relativistic conservation laws,

$$E_\mathrm{i} = E_\mathrm{f} : \qquad\qquad E + m_\mathrm{e}c^2 = E' + E_\mathrm{e}$$
$$p_{x,\mathrm{i}} = p_{x,\mathrm{f}} : \qquad\qquad p = p_\mathrm{e} \cos\varphi + p' \cos\theta$$
$$p_{y,\mathrm{i}} = p_{y,\mathrm{f}} : \qquad\qquad 0 = p_\mathrm{e} \sin\varphi + p' \sin\theta$$

This provides three equations with four unknowns (θ, φ, E_e, E'; p_e and p' are not independent unknowns) that cannot be solved uniquely. Any two of them can be eliminated by solving the equations simultaneously. Choosing to measure the energy and direction of the scattered photons, eliminates E_e and φ. The latter is eliminated by first rewriting the momentum equations as

$$p_\mathrm{e} \cos\varphi = p - p' \cos\theta \qquad \text{and} \qquad p_\mathrm{e} \sin\varphi = p' \sin\theta$$

Squaring and adding yields
$$p_\mathrm{e}^2 = p^2 - 2pp' \cos\theta + p'^2$$
The relativistic relationship between energy and momentum is

$$E_\mathrm{e}^2 = c^2 p_\mathrm{e}^2 + m_\mathrm{e}^2 c^4$$

Substituting this into the equation for E_e obtained by the conservation of energy and for p_e^2 from the above equation yields

$$(E + m_\mathrm{e}c^2 - E')^2 = c^2(p^2 - 2pp' \cos\theta + p'^2) + m_\mathrm{e}^2 c^4$$

Rewriting,
$$\frac{1}{E'} - \frac{1}{E} = \frac{1}{m_\mathrm{e}c^2}(1 - \cos\theta)$$

In terms of wavelengths,
$$\lambda' - \lambda = \frac{h}{m_\mathrm{e}c}(1 - \cos\theta)$$

where λ is the wavelength of the incident photon and λ' is that of the scattered photon, and the quantity
$$\lambda_c = \frac{h}{m_\mathrm{e}c} \approx 0.002\,42\,\mathrm{nm} \qquad\qquad \text{(Compton wavelength of the electron)}$$

is the *Compton wavelength of the electron*. It should be noted that this is a *change* of wavelength rather than a wavelength in and of itself.

The above equations give the change in energy or wavelength as a function of the *scattering angle* θ. As the quantity on the right side is always positive, E' is always less than E, so the scattered photon has less energy than the original incident photon; the difference is simply the kinetic energy given to the electron, $E_\mathrm{e} - m_\mathrm{e}c^2$. Similarly, λ' is greater than λ, meaning that the wavelength of the scattered photon is always longer than that of the incident photon, the change in wavelength ranging from 0 at $\theta = 0°$ to $2\lambda_c$ at $\theta = 180°$. These two descriptions are of course equivalent, the choice of which to use being simply a matter of convenience.

Using
$$E_\mathrm{e} = K_\mathrm{e} + m_\mathrm{e}c^2$$

conservation of energy can also be written as

$$E = m_\mathrm{e}c^2 = E' + K_\mathrm{e} + m_\mathrm{e}c^2$$

Solving for K_e yields

$$K_e = E - E'$$

That is, the kinetic energy of the electron is equal to the difference in the energies of the initial and final photon energies.

The direction of the electron's motion can be found by dividing the momentum relationships:

$$\tan \varphi = \frac{p_e \sin \varphi}{p_e \cos \varphi} = \frac{p' \sin \theta}{p - p' \cos \theta} = \frac{E' \sin \theta}{E - E' \cos \theta}$$

the final result coming from

$$p = \frac{E}{c} \qquad \text{and} \qquad p' = \frac{E'}{c}$$

3.5   Other Photon Processes

Interactions of Photons with Atoms

Electromagnetic radiation is emitted from atoms in discrete amounts characterized by on or more photons. When a photon of energy E is emitted, the atom loses that amount of energy. Consider an atom at rest with initial energy E_i that emits a photon of energy E. Following this, the energy left with the atom is E_f, taken to be that associated with its internal structure. Due to the conservation of momentum, the final atom must have a momentum equal and opposite to that of the emitted photon, giving it a "recoil" kinetic energy K (which is typically quite small). Conservation of energy then gives

$$E_i = E_f + K + E \qquad \text{or} \qquad E = (E_i - E_f) - K$$

The energy of the emitted photon is equal to that lost by the atom plus the negligible recoil kinetic energy of the atom.

An atom can also *absorb* a photon of energy E. If the atom begins at rest, it must acquire a small recoil energy to conserve momentum. Conservation of energy now gives

$$E_i + E = E_f + K \qquad \text{or} \qquad E_f - E_i = E - K$$

The energy available to add to the atom's internal energy is that of the photon minus the typically negligible recoil energy.

Photon emission and absorption provide much insight as to the internal structure of atoms.

Bremsstrahlung and X-ray Production

When a charge, such as an electron, undergoes an acceleration (or deceleration), it radiates electromagnetic energy. According to the quantum interpretation, it emits photons. Suppose there is a beam of electrons that has been accelerated through potential difference ΔV, so that the electrons have a loss of potential energy of $-e\Delta V$, making them gain a kinetic energy of $K = e\Delta V$. When the electrons strike a target, they are slowed and eventually come to a rest, as they collide with the atoms of the target material. In such a collision, momentum is transferred to the atom, the electron slowing down and photons being emitted. The recoil kinetic energy of the atom is negligible (due

to the relative mass of the atom) and can be neglected. If the electron has kinetic energy K prior to the collision and leaves with kinetic energy $K' < K$, then the energy of the photon is

$$hf = \frac{hc}{\lambda} = K - K'$$

The amount of energy lost, and therefore the energy and wavelength of the emitted photon, are not uniquely determined, as the only known energy is K. An electron usually makes many collisions, therefore emitting many different photons, before it is brought to rest. The photons range from very small energies (large wavelengths) corresponding to small losses of kinetic energy to a maximum photon energy of hf_{max} equal to K, corresponding to an electron losing all of its kinetic energy K in a single collision ($K' = 0$). The smallest emitted wavelength λ_{min} is therefore determined by the maximum possible energy loss:

$$\lambda_{\text{min}} = \frac{hc}{K} = \frac{hc}{e\Delta V}$$

For typical accelerating voltages in the range of $10 \times 10^5\,\text{V}$, λ_{min} is on the order of a few tenths of a nanometer, corresponding to the X-ray region of the spectrum. This *continuous* distribution of X-rays (which is different from the *discrete* distribution of X-ray energies emitted in atomic transitions) is called *bremsstrahlung*, which is German for braking/decelerating radiation. Symbolically the bremsstrahlung process can be written as

$$\text{electron} \rightarrow \text{electron} + \text{photon} \qquad \text{(bremsstrahlung process)}$$

The reverse process is simply the photoelectric effect, which is

$$\text{electron} + \text{photon} \rightarrow \text{electron} \qquad \text{(photoelectric effect)}$$

Neither of these processes occur for free electrons. In both cases, there must be a heavy atom to account for the recoil momentum.

Pair Production and Annihilation

When a photon encounters an atom, *pair production* may occur, in which a photon loses all of its energy, in the process creating an electron and a positron. (A positron is an anti-electron, having an identical mass but a positive charge.) This is an example of the creation of rest energy. The electron did not exist prior to the interaction between the photon an the atom. (It was *not* part of the atom.) The photon energy is converted into the relativistic total energies E_+ and E_- of the positron and electron:

$$hf = E_+ + E_- = (m_e c^2 + K_+) + (m_e c^2 + K_-) \qquad \text{(pair production)}$$

As K_+ and K_- are always positive, the photon's energy must be at least $2m_e c^2 = 1.02\,\text{MeV}$ in order for pair production to occur. Such high-energy photons are in the region of *nuclear gamma rays*. Symbolically, this pair production can be represented as

$$\text{photon} \rightarrow \text{electron} + \text{positron} \qquad \text{(pair production)}$$

This process, like bremsstrahlung, can only occur in the presence of a nearby atom to supply the necessary recoil momentum.

The reverse process, *electron-positron annihilation*, can also occur for free electrons and positrons

so long as at least to photons are created (as momentum must be conserved, so the momenta of the photons must cancel). In this process, the electron and positron are replaced by two photons. Symbolically,

$$\text{electron} + \text{positron} \to \text{photon} + \text{photon} \qquad \text{(electron-positron annihilation)}$$

Conservation of energy mandates that

$$(m_e c^2 + K_+) + (m_e c^2 + K_-) = E_1 + E_2$$

where E_1 and E_2 are the energies of the resultant photons. The kinetic energies K_+ and K_- are usually negligibly small, so the electron and positron can be treated as essentially at rest. The conservation of momentum then means that the two photons must have equal and opposite momenta, so their energies must be equal. The energies must be $0.511\,\text{MeV} = m_e c^2$.

Chapter 4

The Wavelike Properties of Particles

4.1   De Broglie's Hypothesis

De Broglie hypothesized that the wave-particle duality of light applies also to material objects in general. Lacking any experimental evidence, he suggested that for any material particle with momentum p, there is a *de Broglie wavelength*

$$\lambda = \frac{h}{p} \qquad \text{(de Broglie wavelength)}$$

Whenever a particle moves, there is an associated de Broglie wave, which shows itself *when a wave-type experiment (such as diffraction) is performed on it.*

4.3   Uncertainty Relationships for Classical Waves

In quantum mechanics, the amplitude of the de Broglie wave yields information regarding the particle's location.

For a quantum particle, the better its momentum (or wavelength) is known, the less is known about its location.

Classical waves also show this effect. All real waves can be represented as *wave packets*, disturbances localized to a finite region of space.

Consider a very small wave packet. The disturbance is localized to a region of length Δx. In attempting to measure the wavelength of this wave packet by placing a measuring stick along the wave, it is difficult to determine where exactly the wave begins and ends. The measurement is therefore subject to a small *uncertainty* $\Delta \lambda$. This uncertainty can be represented as a fraction ε of the wavelength λ such that

$$\Delta \lambda \sim \varepsilon \lambda$$

The fraction ε is less than 1 but likely greater than 0.01, so it may be estimated that

$$\varepsilon \sim 0.1$$

within an order of magnitude. (The $\sim$ symbol is used to indicate a rough order-of-magnitude estimate.) That is, the uncertainty in the measurement is roughly a tenth of the wavelength.

The size of this wave disturbance is approximately one wavelength, so

$$\Delta x \approx \lambda$$

Multiplying this by the uncertainty expression for wavelength yields

$$\Delta x \Delta \lambda \sim \varepsilon \lambda^2 \qquad \text{(classical position-wavelength uncertainty relationship)}$$

For a given wavelength, the smaller the size of the wave packet, the greater the uncertainty in the measurement of its wavelength.

The Frequency-Time Uncertainty Relationship

Consider the measurement of a period rather than the wavelength of the wave comprising the wave packet. Suppose there is a timing device used to measure the wave packet's duration. The wave disturbance is then a function of time rather than location. The "size" of the wave packet is not its duration in time, which is approximately one period T, so

$$\Delta t \approx T$$

As before, it is assumed that the uncertainty is some small fraction of the period, so

$$\Delta T \sim \varepsilon T$$

This yields
$$\Delta t \Delta T \sim \varepsilon T^2 \qquad \text{(classical time-period uncertainty relationship)}$$

For a wave of a given period, the smaller the wave packet's duration, the larger the uncertainty in the measurement of its period.
As frequency and period are related as

$$f = \frac{1}{T}$$

one might assume that

$$\Delta f = \frac{1}{\Delta T} \qquad \text{(incorrect time-frequency relationship)}$$

This, however, is *not* the case, as this would imply that a small uncertainty in the period implies a very large uncertainty in the frequency, which is the inverse of the expected relationship, where more knowledge of one leads to less uncertainty regarding the other. Beginning with the relationship between frequency and time, taking differentials on both sides yields

$$\mathrm{d}f = -\frac{1}{T^2}\,\mathrm{d}T$$

The infinitesimal differentials can then be converted to finite intervals. As only the magnitude of the measurements is of import, the negative may be dropped, yielding

$$\Delta f = \frac{1}{T^2}\Delta T$$

Combining this result with the time-period uncertainty relationship yields

$$\Delta f \Delta t \sim \varepsilon \qquad \text{(classical time-frequency uncertainty relationship)}$$

The longer the duration of the wave packet, the more precisely its frequency can be measured.

4.4 Heisenberg Uncertainty Relationships

The uncertainty relationships discussed prior apply to *all* waves, so they should also apply to de Broglie waves. The basic de Broglie relationship

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

can be used to relate the uncertainty in the momentum Δp and in the wavelength $\Delta \lambda$ first by taking differentials on both sides to yield

$$dp = -\frac{h}{\lambda^2} \, d\lambda$$

and then substituting finite differences for the differentials and removing the negation:

$$\Delta p = \frac{h}{\lambda^2} \Delta \lambda$$

Combining this equation with the time-wavelength uncertainty relationship yields

$$\Delta p \Delta x \sim \varepsilon h$$

The smaller the size of the wave packet, the greater the uncertainty in its momentum (and by proxy its velocity).

Quantum mechanics has a formal procedure for calculating the uncertainties corresponding to different physical situations. One outcome of these calculations yields the wave packet with the smallest possible value of the product, which is $h/4\pi$, meaning that $\varepsilon = 1/4\pi$. In all other cases, then, the product must be greater than this.

The combination $h/2\pi$ is abbreviated as $\hbar$, called *the reduced Planck's constant*:

$$\hbar = \frac{h}{2\pi} \approx 1.05 \times 10^{-34} \, \text{J s} \approx 6.58 \times 10^{-16} \, \text{eV s} \qquad \text{(reduced Planck's constant)}$$

In terms of this constant, the uncertainty relationship can be rewritten as

$$\Delta x \Delta p_x \geq \frac{1}{2}\hbar \qquad \text{(Heisenberg position-momentum uncertainty relationship)}$$

This is the first of the *Heisenberg uncertainty relationships*. It identifies the limits of *simultaneous* measurements. The more a particle is confined, the more uncertain its momentum is.

As $\hbar/2$ is the minimum value of the product, it is generally acceptable to estimate the uncertainty relationship as

$$\Delta x \Delta p_x \sim \hbar$$

The time-period uncertainty relationship can also be applied to de Broglie waves. If it is assumed that the time-frequency relationship for light can also be applied to particles, then the uncertainty relationship becomes

$$\Delta E = h \Delta f$$

Combining this with the time-frequency uncertainty relationship yields

$$\Delta E \Delta t \sim \varepsilon h$$

The minimum uncertainty wave packet gives $\varepsilon = 1/4\pi$, so

$$\Delta E \Delta t \geq \frac{1}{2}\hbar \qquad \text{(Heisenberg energy-time uncertainty relationship)}$$

The more is known about the time coordinate of a particle, the less is known about its energy. For example, a particle with a very short lifetime will have much uncertainty regarding its rest energy. Inversely, the rest energy of a stable particle (one with infinite lifetime) can be measured with effectively infinite precision.

As with the first Heisenberg uncertainty relationship, the second can be reasonable approximated for most wave packets as

$$\Delta E \Delta t \sim \hbar$$

The Heisenberg uncertainty relationships are mathematical representations of the *Heisenberg uncertainty principal*, which states the following:

1. *It is not possible to make a simultaneous determination of the position and momentum of a particle with unlimited precision.*

2. *It is not possible to make a simultaneous determination of the energy and time coordinate of a particle with unlimited precision.*

These relationships estimate the minimum uncertainty from any experiment. The uncertainty may be worse, but it *may be no better*.

4.5   Wave Packets

A *traveling wave* is one that moves in a direction with uniform speed. As the wave packet moves, individual locations oscillate with the frequency or wavelength that characterizes the packet. The displacement of an individual point at a specific time does not matter so much as the locations in space where the overall wave has a large or small oscillation amplitude.

A wave packet can be build by adding waves together. A pure sinusoidal wave is not terribly helpful in representing a particle; as the wave extends over $\mathbb{R}$, the particle could be found anywhere. The particle being represented by the wave packet should be *localized* to a particular region of space.

The key to building a wave packet involves adding waves of different wavelengths. Waves are represented by

$$A\cos(kx)$$

where k is the wave number

$$k = \frac{2\pi}{\lambda}$$

and A is the amplitude. Adding two waves yields

$$y(x) = A_1 \cos(k_1 x) + A_2 \cos(k_2 x) = A_1 \cos\left(\frac{2\pi x}{\lambda_1}\right) + A_2 \cos\left(\frac{2\pi x}{\lambda_2}\right)$$

Adding two different waves yields a reduced amplitude of the wave packet in some locations. This patten repeats over $\mathbb{R}$, though, so the wave is still not localized.

Adding more waves of a greater range of wavelengths helps restrict the size of the wave packet. The above equation can be rewritten using trigonometric identities as

$$y(x) = 2A\cos\left(\frac{\pi x}{\lambda_1} - \frac{\pi x}{\lambda_2}\right)\cos\left(\frac{\pi x}{\lambda_1} + \frac{\pi x}{\lambda_2}\right) \qquad \text{(combined static wave)}$$

If λ_1 and λ_2 are close (that is, if $\Delta\lambda = \lambda_2 - \lambda_1 \ll \lambda_1, \lambda_2$), this can be approximated as

$$y(x) \approx 2A \cos\left(\frac{\Delta\lambda\pi x}{\lambda_{\text{avg}}^2}\right) \cos\left(\frac{2\pi x}{\lambda_{\text{avg}}^2}\right)$$

where

$$\lambda_{\text{avg}} = \frac{\lambda_1 + \lambda_2}{2} \approx \lambda_1 \approx \lambda_2$$

Any finite combination of waves with discrete wavelengths produces patterns that repeat over $\mathbb{R}$, so adding waves does not work in constructing a finite wave packet. To make a finite wave packet, the first cosine term in the above equation must be replaced with a function that is large in the region where the particle is to be confined but falls to 0 elsewhere. The simplest such function is $1/x$, so a wave packet could be

$$y(x) = \frac{2A}{x} \sin\left(\frac{\Delta\lambda\pi x}{\lambda_0^2}\right) \cos\left(\frac{2\pi x}{\lambda_0}\right)$$

where λ_0 is the central wavelength, replacing λ_{avg}. (The first cosine being replaced with a sine is to prevent an infinite discontinuity at $x = 0$.)

Another function with this property is the *Gaussian* modulating function

$$y(x) = A\mathrm{e}^{-2(\Delta\lambda\pi/\lambda_0^2)^2} \cos\left(\frac{2\pi x}{\lambda_0}\right) \qquad \text{(Gaussian modulating function)}$$

Both of these functions display the characteristic inverse relationship between the arbitrarily defined size of the wave packet Δx and the wavelength range parameter $\Delta\lambda$ used in constructing the wave packet.

Wave packets can also be constructed by adding waves of differing amplitudes and wavelengths, but the wavelengths form a continuous set rather than a discrete one. It is easier to see this working with the wave number rather than the wavelength. Adding discrete waves,

$$y(x) = \sum A_i \cos(k_i x) \qquad \text{(discrete wave packet)}$$

Turning the sum into an integral over a continuous set of wave numbers yields

$$y(x) = \int A(k) \cos(kx)\,\mathrm{d}k \qquad \text{(continuous wave packet)}$$

A better approximation of a wave packet's shape comes from letting $A(k)$ vary according to the Gaussian distribution

$$A(k) = A_0 \mathrm{e}^{-(k-k_0)^2/2(\Delta k)^2} \qquad \text{(Gaussian distribution)}$$

This yields a range of wave numbers with its largest contribution at the central wave number k_0 that falls to 0 for larger or smaller wave numbers with characteristic width Δk. Applying the integral equation to this case, integrating over $\mathbb{R}$, yields

$$y(x) = A_0 \Delta\sqrt{2\pi}\mathrm{e}^{(-\Delta k x)^2/2} \cos(k_0 x)$$

Specifying the distribution of wavelengths enables the construction of a wave packet of any desired shape. A wave packet restricting a particle to a region in space of width Δx will have a distribution of wavelengths characterized by width $\Delta\lambda$. As Δx gets smaller, the spread of $\Delta\lambda$ gets larger. This is consistent with the position-wavelength uncertainty relationship for classical waves.

4.6 The Motion of a Wave Packet

To turn a static wave

$$y(x) = A\cos(kx) \qquad \text{(static wave)}$$

into a traveling wave moving in the positive x direction, kx is replaced by $kx - \omega t$, so the traveling wave is written as

$$y(x,t) = A\cos(kx - \omega t) \qquad \text{(traveling wave)}$$

where ω is the *circular frequency* of the wave:

$$\omega = 2\pi f \qquad \text{(circular frequency)}$$

The combined traveling wave is then represented as

$$y(x,t) = A_1 \cos(k_1 x - \omega_1 t) + A_2 \cos(k_2 x - \omega_2 t)$$

The speed of a wave is related to its frequency and wavelength and its circular frequency and wave number as

$$v = \lambda f = \frac{\omega}{k} \qquad \text{(wave speed)}$$

This is also sometimes called the *phase speed* and represents the speed of a particular phase or component of the wave packet. In general, each individual component may have its own speed. As a result, the shape of the wave packet may change over time.

Trigonometric identities can be used to rewrite the combined traveling wave as

$$y(x,t) = 2A\cos\left(\frac{\Delta k}{2}x - \frac{\Delta\omega}{2}t\right)\cos\left(\frac{k_1 + k_2}{2}x - \frac{\omega_1 + \omega_2}{2}t\right) \qquad \text{(combined traveling wave)}$$

The second term represents the rapid variation of the wave within the envelope determined by the first term. The first term is what dictates the overall shape of the wave, so it is this term that determines its speed of travel as well. For a wave written as

$$y(x,t) = \cos(kx - \omega t)$$

the speed is

$$v = \frac{\omega}{k}$$

For the wave envelope, the speed is

$$v_{\text{group}} = \frac{\Delta\omega/2}{\Delta k/2} = \frac{\Delta\omega}{\Delta k}$$

This is called the *group speed* of the wave packet. It can be quite different from the phase speeds of the component waves. For more complicated situations, the group speed can be generalized as the derivative

$$v_{\text{group}} = \frac{\mathrm{d}\omega}{\mathrm{d}k} \qquad \text{(group speed)}$$

The group speed is dependent on the relationship between frequency and wavelength for the component waves. If the phase speeds of all component waves are the same and is independent of frequency or wavelength (such as light waves in empty space) then the group speed is the same as the phase speed and the wave packet maintains its original shape as it travels. In general, a component wave's propagation of a component wave depends on the properties of the medium. De Broglie waves generally have different phase speeds, so their wave packets expand as they travel.

The Group Speed of de Broglie Waves

Consider a localized particle represented by a group of de Broglie waves. The energy of the particle for each component wave is related to the de Broglie frequency by

$$E = hf = \hbar\omega$$

so

$$\mathrm{d}E = \hbar\,\mathrm{d}\omega$$

The momentum of the particle is similarly related to the wavelength by

$$p = \frac{h}{\lambda} = \hbar k$$

so

$$\mathrm{d}p = \hbar\,\mathrm{d}k$$

The group speed of the de Broglie wave can then be written as

$$v_{\text{group}} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{\mathrm{d}E\,/\hbar}{\mathrm{d}p\,/\hbar} = \frac{\mathrm{d}E}{\mathrm{d}p}$$

For a classical particle with only kinetic energy

$$E = K = p^2/2m$$

this can be found to be

$$\frac{\mathrm{d}E}{\mathrm{d}p} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p^2}{2m}\right) = \frac{p}{m} = v$$

which is simply the particle's velocity. Combining this with the prior equation yields an important result:

$$v_{\text{group}} = v_{\text{particle}}$$

The speed of a particle is equal to the group speed of the corresponding wave packet. The wave packet and particle move together.

The Spreading of a Moving Wave Packet

Consider a wave packet that represents a confined particle at $t = 0$. The initial uncertainty in its position is Δx_0 and that in its momentum is $\Delta p_{x,0}$. It moves in the x direction with a velocity v_x with initial uncertainty

$$\Delta v_{x,0} = \frac{\Delta p_{x,0}}{m}$$

This uncertainty means that its location at time t cannot be precisely known; that is,

$$x = v_x t \qquad \text{with} \qquad v_x = v_{x,0} \pm \Delta v_{x,0}$$

There are therefore two contributions to the uncertainty in the wave packet's location at time t: the initial uncertainty Δx_0 and the additional uncertainty due to the momentum $t\Delta v_{x,0}$ that represents

the spreading of the wave packet. It is assumed that this two contributions add quadratically, like experimental uncertainties, making the total uncertainty

$$\Delta x = \sqrt{(\Delta x_0)^2 + (t\Delta v_{x,0})^2} = \sqrt{(\Delta x_0)^2 + \left(\frac{t\Delta p_{x,0}}{m}\right)^2}$$

Taking

$$\Delta p_{x,0} = \frac{\hbar}{\Delta x_0}$$

from the uncertainty principle yields

$$\Delta x = \sqrt{(\Delta x_0)^2 + \left(\frac{\hbar t}{m\Delta x_0}\right)^2}$$

Making Δx_0 very small means that the second term in the square root makes the wave packet expand rapidly. *The more a wave packet is confined, the more quickly it spreads.* Consider the single-slit experiment: the narrower the slit, the more the waves diverge upon passage through it.

Chapter 5

The Schrödinger Equation

5.2 Confining a Particle

A free particle, one with no forces acting on it, is by definition not confined, so it may be located anywhere. It has a definite wavelength, momentum, and energy. A confined particle, on the other hand, is represented by a wave packet that makes it likely to be found only in a region of space of size Δx. This wave packet is constructed through the addition of sines and cosines to obtain the desired mathematical shape.

Consider an electron moving in one dimension confined by a series of electric fields; specifically, consider a *potential energy well*, within which the potential energy is 0 and outside of which it is $U_0 = eV$. To confine this electron, it should be moving in the well with kinetic energy K that is less than U_0.

onsider an infinitely high potential energy barrier between points A and B. This is a good analog to the prior situation. In this case, penetration into the forbidden region cannot occur. The probability of finding the electron anywhere outside the region in which the potential energy is 0 (including at A and B) is therefore 0, as is the amplitude of the wave function at those locations. In order for the wave function to be continuous, the wave function in the 0 potential section must have values of 0 at the boundaries.

Of all possible waves that may describe this particle, the continuity restricts the options to those that are 0 at the boundaries. Note that the function need not be differentiable, only continuous, at the infinite barriers.

Unlike free particles, which may have any wavelength, *only certain values of the wavelength are allowed*. The de Broglie relationship then means that only certain momenta are allowed and consequently only certain energies. The energy is not continuous, instead being discrete. This is known as *quantization of energy*.

The allowed wavelengths are

$$\lambda_n = \frac{2L}{n}, \quad n \in \mathbb{Z}^+$$

where L is the length of the section of 0 potential. This set of wavelengths is identical to those of the classical problem of standing waves on a string stretched between two points. The de Broglie relationship yields

$$p_n = n\frac{h}{2L}, \quad n \in \mathbb{Z}^+$$

The particle's energy is only kinetic, so

$$E_n = n^2 \frac{h^2}{8mL^2}, \quad n \in \mathbb{Z}^+$$

5.3   The Schrödinger Equation

The differential equation whose solution yields the wave behavior of particles is the *Schrödinger equation*. This cannot be derived from any prior equations or postulates; they are fundamental. For nonrelativistic motion, it provides results that correctly account for atomic and subatomic observations.

The form of the Schrödinger equation can be justified by examining the expected solution for a free particle, which should be a wave whose shape at any particular time, as specified by *wave function* $\psi(x)$, is that of a de Broglie wave, such as

$$\psi(x) = A\sin(kx)$$

Differentiating this yields

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = kA\cos(kx) \qquad\qquad \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -k^2 A\sin(kx) = -k^2\psi(x)$$

Note that the second derivative can be written in terms of the original function. With the kinetic energy

$$K = \frac{p^2}{2m} = \frac{(h/\lambda)^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

this can be rewritten as

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -k^2\psi(x) = -\frac{2m}{\hbar^2}K\psi(x) = -\frac{2m}{\hbar^2}(E - U)\psi(x)$$

where $E = K + U$ is the particle's nonrelativistic total energy. For a free particle, $U = 0$, so $E = K$, but as a free particle is being used to find a particular solution to extend to more general cases with potential energy $U(x)$, the equation becomes

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + U(x)\psi(x) = E\psi(x) \qquad\qquad \text{(time-independent Schrödinger equation)}$$

This is the *time-independent Schrödinger equation* for one-dimensional motion.

The solution to the above equation gives the shape of the wave at time $t = 0$. The description of a one-dimensional *traveling* wave must involve both x and t. This is represented by function $\Psi(x, t)$:

$$\Psi(x, t) = \psi(x)\mathrm{e}^{-i\omega t} \qquad\qquad \text{(general time-dependent wave function)}$$

The time dependence is given by the complex exponential term $\mathrm{e}^{-i\omega t}$ where

$$\omega = \frac{E}{\hbar}$$

It is assumed that the potential energy $U(x)$ is known, the goal being to solve for the wave function $\psi(x)$ and energy E *for that particular potential energy*. This is a general example of an *eigenvalue* problem; it is possible to obtain solutions to the equation only for particular values of E, known as the *energy eigenvalues*.

The general procedure for solving the Schrödinger equation is as follows:

1. Write the one-dimensional time-independent Schrödinger equation with the appropriate $U(x)$. If it changes discontinuously ($U(x)$ may be discontinuous but $\psi(x)$ may not), different equation may need to be written for different regions of space.

2. Find a solution $\psi(x)$ to the differential equation.

3. Apply boundary conditions to eliminate extraneous solutions and identify constants. If the potential energy changes discontinuously, apply the conditions on $\psi(x)$ (and usually $\mathrm{d}\psi/\mathrm{d}x$ at the boundary between regions.

As the Schrödinger equation is linear, any constant multiple of a solution is a solution in and of itself.

Probabilities and Normalization

The original goal for solving the Schrödinger equation is to obtain the wave properties of the particle. *The wave function $\psi(x)$ is a wave whose squared absolute amplitude gives the probability of finding the particle in a given region of space.*
Let $P(x)$ be the *probability density* (probability per unit length in one dimension). According to the Schrödinger procedure, then according to the interpretation of the wave function,

$$P(x)\,\mathrm{d}x = |\psi(x)|^2\,\mathrm{d}x$$

$|\psi(x)|^2\,\mathrm{d}x$ gives the probability of finding the particle in interval $\mathrm{d}x$ and x (the probability that it lies between x and $x + \mathrm{d}x$).[1] As the wave function may be complex, it is necessary to square its absolute value to ensure that the probability is a positive real number.
The squared magnitude of the general time-dependent wave function is

$$|\Psi(x,t)|^2 = |\psi(x)|^2|\mathrm{e}^{-i\omega t}|^2 = |\psi(x)|^2$$

as applying Euler's formula yields

$$|\mathrm{e}^{-i\omega t}|^2 = |\cos(-\omega t) + i\sin(-\omega(t))|^2 = |\cos(\omega t) - i\sin(\omega t)|^2 = \cos^2(\omega t) + \sin^2(\omega t) = 1$$

For this reason, the probability density associated with a solution to the Schrödinger equation (for any energy eigenvalue) is independent of time. Such quantum states are called *stationary states.*
This interpretation of $|\psi(x)|^2$ clarifies the continuity condition of $\psi(x)$. The probability cannot change discontinuously.
This interpretation of $\psi(x)$ enables the completion of the procedure for solving the Schrödinger equation:

4. For a wave function describing a single particle, the probability summed over all locations must be 1; that is, the particle must be *somewhere* between $-\infty$ and ∞:

$$\int_{-\infty}^{\infty} |\psi(x)|^2\,\mathrm{d}x = 1$$

    The constant multiple of the solution used must fulfill this condition. A wave function with its multiplicative constant chosen in this way is said to be *normalized*, the above equation being known as the *normalization condition.*

---

[1]The probability of finding the particle at a given point x is 0, as a single point is a mathematical abstraction without physical dimension.

5. As the Schrödinger equation represents a probability, any solution that goes to infinity must be discarded (its multiplicative constant set to 0). If the solution is

$$\psi(x) = Ae^{kx} + Be^{-kx}$$

for the *entire* region $x > 0$, then A must be 0 for the solution to be physically meaningful. If this solution is valid for the *entire* region $x < 0$, though, then B must be 0. If the solution is valid only over a small range (say $0 < x < L$, then neither can be 0).

6. Suppose the interval between two points x_1 and x_1 is divided into infinitesimal intervals of width dx. The total probability of the particle being located within the interval, notated $P(x_1 : x_2)$, is

$$P(x_1 : x_2) = \int_{x_1}^{x_2} P(x)\,dx = \int_{x_1}^{x_2} |\psi(x)|^2\,dx$$

If the wave function has been properly normalized, this will always yield a probability between 0 and 1.

7. The outcome of a single measurement cannot be deterministically predicted. Instead, the *average* outcome of a large number of measurements an be predicted:

$$x_{\text{avg}} = \frac{\sum n_i x_i}{\sum n_i}$$

where n_i is the number of times each x_i is observed, proportional to the probability $P(x_i)\,dx$ of finding that particle in the interval dx at i. Making this substitution and replacing the sums with integrals yields

$$x_{\text{avg}} = \frac{\int_{-\infty}^{\infty} P(x)x\,dx}{\int_{-\infty}^{\infty} P(x)\,dx} = \int_{-\infty}^{\infty} |\psi(x)|^2 x\,dx$$

as the denominator is 1 due to the normalization. By analogy, the average value of any function f of x can be found as

$$[f(x)]_{\text{avg}} = \int_{-\infty}^{\infty} P(x)f(x)\,dx = \int_{-\infty}^{\infty} |\psi(x)|^2 f(x)\,dx$$

Average values calculated according to the above equations are *expectation values*.

5.4   Applications of the Schrödinger Equation

Solutions for Constant Potential Energy

Consider the Schrödinger equation for the case with constant potential energy U_0, which is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U_0\psi(x) = E\psi(x)$$

or (assuming that $E > U_0$)

$$\frac{d^2\psi}{dx^2} = -k^2\psi(x) \qquad \text{where} \qquad k = \sqrt{\frac{2m(E - U_0)}{\hbar^2}}$$

The parameter k is simply wave number $2\pi/\lambda$. The auxiliary equation for this differential equation is

$$m^2 + k^2 = 0$$

which has solutions $m = \pm ki$, making the solution to the differential equation

$$\psi(x) = A\sin(kx) + B\cos(kx)$$

The values of the constants can be determined by applying the continuity and normalization requirements.

To analyze the penetration of a particle into a forbidden region, the energy E of the energy should be smaller than the potential energy U_0. For this case, the differential equation becomes

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = k'^2\psi(x) \quad \text{where} \quad k' = \sqrt{\frac{2m(U_0 - E)}{\hbar^2}}$$

The auxiliary equation for this is

$$m^2 - k'^2 = 1$$

which has solutions $m = \pm k$, making the solution to the differential equation

$$\psi(x) = A\mathrm{e}^{k'x} + B\mathrm{e}^{-k'x}$$

The Free Particle

The net force on a free particle is 0, making its potential energy constant. Any value for this constant can be used, so for convenience, $U_0 = 0$. The solution is therefore

$$\psi(x) = A\sin(kx) + B\cos(kx)$$

The particle's energy is

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

No restrictions are placed on k, meaning that it (and by proxy E) can take on any value, making it *not* quantized. Note that this kinetic energy is that of a particle with momentum

$$p = \hbar k = \frac{h}{\lambda}$$

This is as expected, as the free particle can be replaced by any de Broglie wave.

The normalization integral cannot be evaluated over $\mathbb{R}$, meaning that the probabilities cannot be determined from the wave function.

Writing the wave function in terms of complex exponentials,

$$\psi(x) = A\left(\frac{\mathrm{e}^{ikx} - \mathrm{e}^{-ikx}}{2i}\right) + B\left(\frac{\mathrm{e}^{ikx} + \mathrm{e}^{-ikx}}{2}\right) = A'\mathrm{e}^{ikx} + B'\mathrm{e}^{ikx}$$

where

$$A' = \frac{A}{2i} + \frac{B}{2} \quad \text{and} \quad B' = -\frac{A}{2i} + \frac{B}{2}$$

The time-dependent wave function can then be written

$$\Psi(x,t) = (A'\mathrm{e}^{ikx} + B'\mathrm{e}^{-ikx})\mathrm{e}^{-i\omega t} = A'\mathrm{e}^{i(kx-\omega t)} + B'\mathrm{e}^{-i(kx+\omega t)}$$

The first term's dependence on $kx - \omega t$ means that it represents a wave moving in the positive x direction with amplitude A', and the second term's on $kx + \omega t$ makes it a wave moving in the negative x direction with magnitude B'.

To make this wave represent a beam of particles moving in the positive x direction, B' must be 0. The probability density associated with this wave is then

$$P(x) = |\psi(x)|^2 = |A'|^2 e^{ikx} e^{-ikx} = |A'|^2$$

The probability density is constant, meaning that the particles are equally likely to be at any given point on the x axis. This is consistent with the de Broglie wave: a wave of a precisely defined wavelength extends over $\mathbb{R}$, thus giving a completely unlocalized particle.

Infinite Potential Energy Well

Consider a particle trapped between $x = 0$ and $x = L$ by barriers of infinitely high potential energy. The particle moves freely in this region, making elastic collisions with the perfectly rigid barriers. (This problem is sometimes called "a particle in a box".) For now it is assumed that the particle only moves in a single spatial dimension. The potential energy can be expressed as

$$U(x) = \begin{cases} 0 & x \in [0, L] \\ \infty & x \notin [0, L] \end{cases}$$

(Between 0 and L, U may be any constant value, but it is set to 0 for convenience.) As the potential energy differs within and without the well, separate solutions are required for each region.

Considering the time-independent Schrödinger equation outside the well, it can be seen that the only way for $P(x)$ to be meaningful as $U \to \infty$ is for ψ to be 0, so that $U\psi$ will remain finite. Alternatively, it can be reasoned that should the barriers be perfectly rigid, the particle must always be within the well, so the probability of finding it outside must be 0, meaning that ψ must be 0. This yields

$$\psi(x) = 0 \qquad x \notin [0, L]$$

The Schrödinger equation for within the well is

$$-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 \psi}{\mathrm{d}x^2} = E\psi(x)$$

which is simply that for constant potential with $U_0 = 0$, making the solution

$$\psi(x) = A\sin(kx) + B\cos(kx), \qquad x \in [0, L]$$

where

$$k = \sqrt{\frac{2mE}{\hbar^2}}$$

As the wave function must be continuous, both cases must have the same values at the boundaries. At $x = 0$,

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

Applying the boundary condition $\psi(0) = 0$ means that $B = 0$. At $x = L$,

$$\psi(x) = A\sin(kL) + B\cos(kL) = A\sin(kL)$$

Applying the second boundary condition $\psi(L) = 0$ (and placing the restriction that $A \neq 0$ so that the solution has some meaning) yields

$$kL = n\pi, \quad n \in \mathbb{Z}^+$$

Using the definition of the wave number yields

$$\lambda_n = \frac{2L}{n}, \quad n \in \mathbb{Z}^+$$

The is identical to the result obtained for standing waves in a string of length L fixed at both ends. *Thus the solution to the Schrödinger equation for a particle trapped in a linear region of fixed length is a series of standing de Broglie waves.*

As only certain values of k are allowed, it can be seen that *energy is quantized*:

$$E_n = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2} = \frac{h^2 n^2}{8mL^2}, \quad n \in \mathbb{Z}^+$$

Let

$$E_0 = \frac{\hbar^2 \pi^2}{2mL^2} = \frac{h^2}{8mL^2}$$

This yields

$$E_n = n^2 E_0, \quad n \in \mathbb{Z}^+$$

The lowest energy state ($n = 1$) is the *ground state* while higher energies ($n > 1$) are *excited states*. As the only form of energy in this scenario is kinetic, the particle's velocity is quantized. Classically, a particle given any initial velocity would move back and forth at the same speed, but in the quantum case, only certain values yield sustained states of motion; these special conditions represent the "stationary states".

The particle can move between energy states by absorbing energy (to move up) or releasing it (to move down). A similar effect occurs for electrons within atoms, in which the absorbed or released energy is typically a photon.

Applying the normalization condition,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = \int_0^L A^2 \sin^2\left(\frac{n\pi x}{L}\right) dx$$

which yields

$$A = \sqrt{\frac{2}{L}}$$

making the complete wave function for $[0, L]$

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \quad , \quad n \in \mathbb{Z}^+$$

Note that n corresponds directly to the number of peaks in the probability density.

The normalization condition need not be applied simply to find probabilities:

$$P(x_1 : x_2) = \int_{x_1}^{x_2} |\psi_n(x)|^2 \, dx = \frac{\int_{x_1}^{x_2} \sin^2(n\pi x/L)}{\int_0^L \sin^2(n\pi x/L)}$$

This simply the proportion of area of the density function contained within the interval.

Finite Potential Energy Well

Consider a potential energy

$$U(x) = \begin{cases} 0 & x \in [0, L] \\ U_0 & x \notin [0, L] \end{cases}$$

In order for the particle to be confined to the well, its energy must be less than U_0.

Over $[0, L]$, the Schrödinger equation and its solution remains the same as that for the infinite potential case:

$$\psi(x) = A\sin(kx) + B\cos(kx) \qquad x \in [0, L]$$

The previously found values of the constants are not applicable, though, as the boundary conditions are now different. For $x < 0$, the solution must be of the form

$$\psi(x) = Ce^{k'x} + De^{-k'x} \qquad x < 0$$

due to the fact that $E < U_0$. As $x \to -\infty$, the second term goes to ∞, meaning that D must be 0:

$$\psi(x) = Ce^{k'x} \qquad x < 0$$

The form of the solution for $x > L$ must be of the same form for the same reason:

$$\psi(x) = Fe^{k'x} + Ge^{-k'x} \qquad x > 0$$

As $x \to \infty$, the first term goes to ∞, meaning that F must be 0:

$$\psi(x) = Ge^{-k'x} \qquad x > L$$

These 3 equations, the 2 boundary conditions, and the normalization condition provides 5 equations, which must be used to solve for the unknown constants A, B, D, and G, as well as the energy E.

The resulting energy values cannot be solved for directly but must instead be found numerically by solving a transcendental equation.

Two-Dimensional Infinite Potential Energy Well

The Schrödinger equation can be extended to 2 spatial dimensions as

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U\psi = E\psi \qquad \text{(2D time-independent Schrödinger equation)}$$

where

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

The potential energy is

$$U(x, y) = \begin{cases} 0 & x, y \in [0, L] \\ \infty & x, y \notin [0, L] \end{cases}$$

Only *separable* solutions to the Schrödinger are considered; that is,

$$\psi(x, y) = f(x)g(y)$$

where

$$f(x) = A\sin(k_x x) + B\cos(k_x y) \qquad \text{and} \qquad g(y) = A\sin(k_y x) + B\cos(k_y y)$$

5.5 The Simple Harmonic Oscillator

The classical oscillator is a mass m attached to a spring of spring constant k. The force exerted by the spring on the mass is

$$F = -kx \qquad \text{(spring force)}$$

where x is the displacement from the equilibrium. Applying Newton's laws, it can be shown that the oscillator has circular/angular frequency

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad \text{(angular frequency)}$$

and period

$$T = \frac{2\pi}{\omega_0} \qquad \text{(period)}$$

The maximum displacement of the mass from its equilibrium position is the amplitude of the oscillation x_0. The kinetic energy is maximized at $x = 0$ while it is 0 at the *turning points* $\pm x_0$. At these turning points, the oscillator comes to a rest for an instant before reversing its direction of motion. The motion is of course confined to $[-x_0, x_0]$.

Any system in a smoothly varying potential energy well near its minimum acts approximately as a simple harmonic oscillator. A force $F = -kx$ has an associated potential energy

$$U = \frac{1}{2}kx^2$$

yielding the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{1}{2}kx^2\psi = E\psi$$

There are no boundaries between regions of potential energy, so the only conditions are that the wave function must fall to 0 as x approaches $\pm\infty$. The simplest function that satisfies these conditions is

$$\psi(x) = A\mathrm{e}^{-ax^2}$$

The constants a and energy E can be found by substitution this into the Schrödinger equation. Differentiating yields

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = -2ax\left(A\mathrm{e}^{-ax^2}\right) = -2ax\psi$$

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -2a\psi - 2ax(-2ax)\psi = (-2a + 4a^2x^2)\psi$$

Substituting this into the Schrödinger equation and cancelling common factor ψ yields

$$\frac{\hbar^2 a}{m} - \frac{2a^2\hbar^2}{m}x^2 + \frac{1}{2}kx^2 = E$$

The desired solution must be valid for any value of x, so the coefficients of x^2 must cancel and the remaining constants must be equal:

$$-\frac{2a^2\hbar^2}{m} + \frac{1}{2} = 0 \qquad \text{and} \qquad \frac{\hbar^2 a}{m} = E$$

which yield

$$a = \frac{\sqrt{km}}{2\hbar} \qquad \text{and} \qquad E = \frac{1}{2}\hbar\sqrt{\frac{k}{m}}$$

The energy can be rewritten in terms of the classical frequency ω_0 as

$$E = \frac{1}{2}\hbar\omega_0$$

making the wave function

$$\psi(x) = Ae^{-x^2\sqrt{km}/2\hbar}$$

Applying the normalization condition,

$$1 = \int_{-\infty}^{\infty} |A|^2 e^{-2ax^2}\,\mathrm{d}x$$

Letting $z = \sqrt{2a}x$,

$$1 = \frac{|A|^2}{\sqrt{2a}}\int_{-\infty}^{\infty} e^{-z^2}\,\mathrm{d}z$$

Letting I be equal to the integral and using the fact that a change in variables does not affect the value of the integral,

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,\mathrm{d}y\,\mathrm{d}x$$

Changing to polar coordinates,

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\int_0^{\infty} e^{-r^2} r\,\mathrm{d}r$$

Letting $u = r^2$,

$$I^2 = \pi\int_0^{\pi}\int_0^{\infty} e^{-u}\,\mathrm{d}u = \pi[-e^{-u}]_0^{\infty} = \pi[0-(-1)] = \pi$$

making the final value of the original integral

$$I = \sqrt{\pi}$$

Substituting this back into the expression for the normalization condition,

$$1 = \frac{|A|^2}{\sqrt{2a}}\sqrt{\pi}$$

Solving for A,

$$A = \left(\frac{2a}{\pi}\right)^{1/4} = \left(\frac{2\sqrt{km}}{2\hbar\pi}\right)^{1/4} = \left(\frac{m\omega_0}{\hbar\pi}\right)^{1/4}$$

making the ground state wave function

$$\psi(x) = \left(\frac{m\omega_0}{\hbar\pi}\right)^{1/4} e^{-x^2\sqrt{km}/2\hbar} \qquad\qquad \text{(simple harmonic oscillator)}$$

The general solution is of the form

$$\psi_n(x) = Af_n(x)e^{-ax^2}$$

where $f_n(x)$ is a polynomial of degree n. The corresponding energies are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0, \quad n \in \mathbb{Z}^+$$

Chapter 6

The Rutherford-Bohr Model of the Atom

6.1 Basic Properties of Atoms

1. *Atoms are very small*, on the scale of a tenth of a nanometer in radius. Any effort to "see" an atom using visible light $\lambda = 500\,\text{nm}$ are thus hopeless owing to diffraction effects. An estimate of the *maximum* size can be made, though, by considering the volume of a mole of an element (calculated from its density) and dividing it by Avogadro's number.

2. *Atoms are stable*; that is, they do not spontaneously break into smaller pieces or collapse. Therefore, the internal forces holding atoms together must be in equilibrium. This means that said forces must be opposed in some way, as otherwise, the atoms would collapse.

3. *Atoms contain negatively charged electrons, but are electrically neutral.* If an atom or collection of atoms is disturbed with sufficient force, electrons are emitted. This is observed in the Compton and photoelectric effects. While electrons are emitted from atoms' nuclei in certain radioactive decay processes, they don't "exist" in those nuclei, instead being manufactured by some process. The uncertainty principle forbids emitted electrons of the energies observed from existing in the nucleus. In a volume as large as an atom, though, the uncertainty principle places no such restriction on the existence of electrons. It can also be observed that bulk matter is electrically neutral and it can be reasonably extrapolated that this is a property of atoms themselves.

4. *Atoms absorb and emit electromagnetic radiation.* The observation of these emissions is crucial to understanding the atom. It should be noted that emitted and absorbed wavelengths may differ.

6.2 Scattering Experiments and the Thomson Model

The Thomson model incorporated already known properties of atoms, such as their size, mass, number of electrons, and electric neutrality. According to this model, an atom contains Z electrons within a sphere of positive charge Ze. As the electrons don't contribute significantly to the mass of the atom, the mass of the sphere is essentially that of the atom. The radius of the sphere R and that of the atom are the same.

6.3 The Rutherford Nuclear Atom

Rutherford concluded that the most likely way for an alpha particle ($m = 4\,\mathrm{u}$) to be deflected through large angles is for it to undergo a single collision with a more massive object. He therefore proposed that an atom's mass and charge are concentrated at a central *nucleus*. The projectile, of charge ze, experiences a deflective Coulomb force due to the nucleus of charge Ze:

$$F = \frac{1}{4\pi\varepsilon_0}\frac{(ze)(Ze)}{r^2} = \frac{zZe^2}{4\pi\varepsilon_0 r^2} \qquad\text{(Rutherford scattering)}$$

Due to their relatively small masses, the atomic electrons have negligible impact on the path of the particle. It is also assumed that the nucleus is far more massive than the particle, meaning that it does not move through the interaction; as no kinetic energy is given to the nucleus, that of the particle remains unchanged.

Each impact parameter b has a certain scattering angle θ. The particle's path can be shown to be hyperbolic; in polar coordinates r and φ,

$$\frac{1}{r} = \frac{1}{b}\sin\varphi + \frac{zZe^2}{8\pi\varepsilon_0 b^2 K}(\cos\varphi - 1) \qquad\text{(Rutherford scattering path)}$$

The particle's initial position is $\varphi = 0, r \to \infty$ and its final position is $\varphi = \pi - \theta, r \to \infty$. Using these final coordinates, the above equation can be reduced to

$$b = \frac{zZe^2}{8\pi\varepsilon_0 K}\cot\left(\frac{1}{2}\theta\right) = \frac{zZ}{2K}\frac{e^2}{4\pi\varepsilon_0}\cot\left(\frac{1}{2}\theta\right) \qquad\text{(Rutherford scattering path)}$$

Note that

$$\frac{e^2}{4\pi\varepsilon_0} = 1.44\,\mathrm{eV\,nm}$$

A positively charged projectile slows as it approaches the nucleus, losing kinetic energy but gaining electrostatic potential energy

$$U = \frac{1}{4\pi\varepsilon_0}\frac{zZe^2}{r^2}$$

Kinetic energy is maximized and potential minimized when the radius is minimized. When the projectile is far from the nucleus, U is 0, so

$$E = K = \frac{1}{2}mv^2$$

As the projectile approaches the nucleus, the total energy remains the same, so at distance $r_{\min}$ and speed $v_{\min}$,

$$E = \frac{1}{2}mv_{\min}^2 + \frac{1}{4\pi\varepsilon_0}\frac{zZe^2}{r_{\min}^2} = \frac{1}{2}mv^2$$

In addition to energy, angular momentum is conserved:

$$L = mvb = mv_{\min}r_{\min}$$

which yields

$$v_{\min} = \frac{bv}{r_{\min}}$$

Substituting this into the energy expression,

$$\frac{1}{2}mv^2 = \frac{1}{2}m\left(\frac{b^2v^2}{r_{\min}}\right) + \frac{1}{4\pi\varepsilon_0}\frac{zZe^2}{r_{\min}}$$

Note that K is not 0 unless b is as well, in which case all of the particle's energy becomes potential, the minimum distance to the nucleus being minimized. At this point, the distance from the nucleus is d, the distance of closest approach. This can be found by solving the above equation for $r_{\min}$ when $b = 0$:

$$d = \frac{1}{4\pi\varepsilon_0}\frac{zZe^2}{K} \qquad\qquad \text{(distance of closest approach)}$$

6.4   Line Spectra

The radiation emitted by atoms can be classified into continuous or discrete/line spectra. In the former case, all wavelengths within some interval are emitted, but this is not the case for line spectra.

All groupings of lines in the hydrogen spectrum can be fit by

$$\lambda = \lambda_{\text{limit}}\frac{n^2}{n^2 - n_0^2}, \quad n = n_0 + 1, n_0 + 2, \ldots \qquad\qquad \text{(hydrogen spectra)}$$

where

$$\lambda_{\text{limit}} = 91.13n_0^2$$

in nm. $n_0 = 1$ is the Lyman series, 2 is Balmer, 3 is Paschen, 4 is Brackett, and 5 is Pfund.

6.5   The Bohr Model

Consider an electron orbiting a hydrogen nucleus. The centripetal force is provided by the attractive Coulomb force between it and the nucleus:

$$F = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r^2} = \frac{mv^2}{r}$$

Manipulating this equation yields the kinetic energy of the electron to be

$$K = \frac{1}{2}mv^2 = \frac{1}{8\pi\varepsilon_0}\frac{e^2}{r}$$

The potential energy of the electron is the Coulomb potential energy

$$U = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r} = -2K$$

The total energy is simply the sum of the kinetic and potential energies:

$$E = K + U = K - 2K = -\frac{1}{8\pi\varepsilon_0}\frac{e^2}{r}$$

Classically, an accelerating charge, such as this electron, must continuously emit electromagnetic radiation, meaning that its total energy would decrease, causing it to spiral towards the nucleus.

Bohr hypothesized that there exist *stationary states* in which an electron may exist without radiating electromagnetic energy. In these states, its angular momentum takes integer multiples of $\hbar$. This is known as the *quantization of angular momentum.*:

$$L = n\hbar, \quad n \in \mathbb{Z}^+ \qquad \text{(quantization of angular momentum)}$$

For a body in circular motion, the momentum and radial vectors are always perpendicular, so

$$L = |\vec{r} \times \vec{p}| = mvr$$

Bohr's postulate is thus

$$mvr = n\hbar, \quad n \in \mathbb{Z}^+ \qquad \text{(Bohr's postulate)}$$

Solving for v and substituting into the equation for kinetic energy yields

$$K = \frac{1}{2}\left(\frac{n\hbar^2}{mr}\right) = \frac{1}{8\pi\varepsilon_0}\frac{e^2}{r}$$

This yields a series of allowed radii:

$$r_n = \frac{4\pi\varepsilon_0\hbar^2}{me^2}n^2, \quad n \in \mathbb{Z}^+$$

This can be rewritten as

$$r_n = a_0 n^2, \quad n \in \mathbb{Z}^+$$

where a_0 is the *Bohr radius*, defined as

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{me^2} \approx 0.0529\,\text{nm} \qquad \text{(Bohr radius)}$$

Substituting the allowed radii into the equation for total energy yields

$$E_n = -\frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2}\frac{1}{n^2} \approx -\frac{13.6\,\text{eV}}{n^2}, \quad n \in \mathbb{Z}^+$$

This means that the electron's energy is *quantized*. E_1 is the *ground state* while the higher states are the *excited states*. (Note that the first excited state is $n = 2$.)

The *excitation energy* of an excited state n is the energy above the ground state:

$$\Delta E = E_n - E_1 \qquad \text{(excitation energy)}$$

This can be regarded as the amount of energy the atom must absorb in order for the electron to make an upward jump.

The magnitude of an electron's energy is sometimes called its *binding energy*. Should the atom absorb an amount of energy equal to the binding energy of an electron, said electron will be removed from the atom, becoming a free electron and making the atom an *ion*. The amount of energy required to remove an electron from an atom is also called the *ionization energy*. The ionization energy of an atom typically indicates the energy required to remove an electron from the ground state. Should an atom absorb sufficient energy to free an electron, the excess energy will become the electron's kinetic energy.

The binding energy can also be regarded as the energy released when an atom is assembled from an electron and nucleus that are initially separated by a large distance. Bringing the electron from a large distance (where $E = 0$) and placing it in orbit in state n results in the change in energy being equal to the negative value of E_n. To compensate for this, energy amounting to $|E_n|$ is released, usually in the form of photons.

The Hydrogen Wavelengths in the Bohr Model

Bohr postulated that electrons emit radiation only when dropping energy levels. The energy of the emitted photon is equal to the change in energy:

$$hf = E_{n_1} - E_{n_2}$$

Using the equation for those energies,

$$f = \frac{me^4}{64\pi^3\varepsilon_0^2\hbar^3}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right)$$

The emitted wavelength is then

$$\lambda = \frac{c}{f} = \frac{64\pi^3\varepsilon_0^2\hbar^3 c}{me^4}\left(\frac{n_1^2 n_2^2}{n_1^2 - n_2^2}\right)$$

which can be rewritten as

$$f = \frac{1}{R_\infty}\left(\frac{n_1^2 n_2^2}{n_1^2 - n_2^2}\right)$$

where R_∞ is the *Rydberg constant*, defined as

$$R_\infty = \frac{me^4}{64\pi^3\varepsilon_0^2\hbar^3 c} \approx 1.097\,373 \times 10^7\,\mathrm{m}^{-1} \qquad\qquad \text{(Rydberg constant)}$$

The Bohr formulas are remarkably consistent with the two longest wavelengths of the Balmer series. They also explain the Ritz combination principle, which states that certain frequencies in the emission spectrum can be summed to yield other frequencies. Consider a transition from n_3 to n_2 and then from n_2 to n_1:

$$f_{n_3 \to n_2} = cR_\infty\left(\frac{1}{n_3^2} - \frac{1}{n_2^2}\right) \qquad\qquad f_{n_2 \to n_1} = cR_\infty\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right)$$

Thus

$$f_{n_3 \to n_2} + f_{n_2 \to n_1} = cR_\infty\left(\frac{1}{n_3^2} - \frac{1}{n_2^2}\right) + cR_\infty\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) = cR_\infty\left(\frac{1}{n_3^2} - \frac{1}{n_1^2}\right)$$

which is simply the frequency of the photon emitted in a direct transition from n_3 to n_1, so

$$f_{n_3 \to n_2} + f_{n_2 \to n_1} = f_{n_3 \to n_1} \qquad\qquad \text{(Ritz transition principle)}$$

As the frequency of the emitted photon is directly proportional to its energy, the Ritz combination principle can be restated in terms of energy: the energy of a photon emitted in a transition that spans multiple states is equal to the sum of the energies of the individual transitions.

The Bohr model also helps explain why atoms don't absorb and emit radiation at all the same wavelengths. Isolated atoms are usually only found in the ground state, as the excited states have lifespans shorter than a nanosecond, after which they decay to the ground state. *The absorption spectrum therefore only contains transitions from the ground state.*

Atoms with $Z > 1$

The Bohr theory can be extended to any atom with a single electron regardless of its nuclear charge Z. For a nucleus of charge Ze, the Coulomb force is

$$F = \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r^2}$$

That is, e^2 is replaced by Ze^2. Making this substitution, the allowed radii and energies become

$$r_n = \frac{4\pi\varepsilon_0\hbar^2}{Ze^2 m}n^2 = \frac{a_0 n^2}{Z} \quad \text{and} \quad E_n = -\frac{m(Ze^2)^2}{32\pi^2\varepsilon_0^2\hbar^2}\frac{1}{n^2} = -(13.6\,\text{eV})\frac{Z^2}{n^2}, \quad n \in \mathbb{Z}^+$$

It is evident that the orbits in higher-Z atoms are closer to the nucleus and have higher (negative) energies; that is, the electron is more tightly bound to the nucleus.

Chapter 7

The Hydrogen Atom in Wave Mechanics

7.1 A One-Dimensional Atom

Quantum mechanics yields a very different view of the atom than the Bohr model. Rather than the electron traveling in a circular orbit about the proton, a fixed radius or orbital plane is not allowed by quantum mechanics. Instead, it describes the electron in terms of a probability density, leading to uncertainty in locating the electron.

In order to analyze the hydrogen atom quantum mechanically, the Schrödinger equation must be solved for the Coulomb potential energy of the proton and electron:

$$U(r) = -\frac{e^2}{4\pi\varepsilon_0 r}$$

Consider a one-dimensional atom, the proton being fixed at the origin and the electron moving along the positive x-axis. In one dimension, the Schrödinger equation for the electron becomes

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - \frac{e^2}{4\pi\varepsilon_0 x}\psi(x) = E(x)$$

For a bound state, the wave function must fall to 0 as x approaches infinity. Additionally, for the second term on the left side to remain finite at the origin, the wave function must be 0 at the origin. The simplest function satisfying both of these requirements is

$$\psi(x) = Ax\mathrm{e}^{-bx}$$

Substituting this trial wave function into the Schrödinger equation yields a solution when

$$b = \frac{me^2}{4\pi\varepsilon_0 \hbar^2} = \frac{1}{a_0}$$

The corresponding energy is

$$E = -\frac{\hbar^2 b^2}{2m} = -\frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2}$$

which is identical to the ground state energy in the Bohr model.

The most probably region for the electron to be in is around $x = a_0$, but there is a nonzero probability of the electron being found energy on the positive x-axis. This differs greatly from the Bohr model, which fixes the distance between the proton and electron at a_0.

Higher excited states have more peaks in the probability density and the region of maximum probability moves further along the x-axis.

7.2   Angular Momentum in the Hydrogen Atom

Angular Momentum of Classical Orbits

Classically, the angular momentum of a particle is

$$\vec{L} = \vec{r} \times \vec{p} \qquad\qquad \text{(classical angular momentum)}$$

where $\vec{r}$ is the position vector and $\vec{p}$ is its linear momentum. The angular momentum is perpendicular to the plane of the orbit. Along with the energy, the angular momentum remains constant in orbit. The total energy of the orbital motion is what determines the average distance from the center of the orbit. For a given total energy, many orbits are possible, from nearly circular to highly elliptical ones. The complete specification of the orbit requires knowledge of the direction of the angular momentum in addition to its magnitude. The direction is what determines the plane of the orbit. To completely describe the angular momentum vector, three numbers are required, regardless of whether Cartesian or spherical coordinates are being used.

Angular Momentum in Quantum Mechanics

Quantum mechanics has quite a different view of angular momentum. The angular momentum properties of a three-dimensional wave function are described by two quantum numbers, the first of which is the *angular momentum* or *azimuthal quantum number* ℓ, which determines the length of the angular momentum vector:

$$L = \hbar\sqrt{\ell(\ell + 1)}, \quad \ell \in \mathbb{N}$$

This differs significantly from the Bohr condition, allowing the angular momentum to be 0.
The second number is the *magnetic quantum number* m_ℓ, which determines one component of the angular momentum vector, typically chosen to be the z component:

$$L_z = m_\ell \hbar, \quad m_\ell = 0, \pm 1, \pm 2, \ldots, \pm \ell$$

Note that each value of ℓ has $2\ell + 1$ possible values of m_ℓ.
Unlike the classical angular momentum vector, which must be specified by three numbers, *only two* are required for the quantum angular momentum vector. These two numbers cannot completely identify a vector in three-dimensional space, so they must signify something else.

The polar angle[1] made between $\vec{L}$ and the z-axis is determined by

$$\cos\theta = \frac{L_z}{L} = \frac{m_\ell}{\sqrt{\ell(\ell+1)}}$$

This implies *spatial quantization*, meaning that only certain orientations of the angular momentum vector are allowed. The number of these orientations is equal to the number of possible values of m_ℓ and the magnitudes of their successive z components always differs by $\hbar$. Classical angular momentum, of course, can have any orientation.

7.3   The Hydrogen Atom Wave Functions

To completely describe the electron in the hydrogen atom, a three-dimensional wave function must be obtained. The three dimensional Schrödinger equation in Cartesian coordinates has the form

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x,y,z) + U(x,y,z)\psi(x,y,z) = E\psi(x,y,z)$$

where

$$\nabla^2\psi(x,y,z) = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}$$

To solve a partial differential equation of this form, the variables are separated by writing the function of three variables as the product of three functions of a single variable, such as

$$\psi(x,y,z) = X(x)Y(y)Z(z)$$

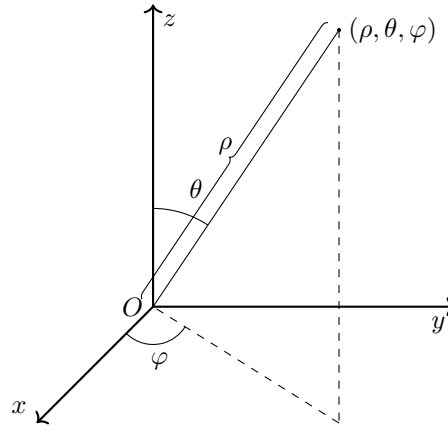The Coulomb potential energy written in Cartesian coordinates is

$$U(x,y,z) = -\frac{e^2}{4\pi\varepsilon_0\sqrt{x^2+y^2+z^2}}$$

---

[1] In mathematics, the convention is for the polar angle to be in the xy-plane and for the azimuthal angle to be from the z-axis, while the conventions for physics swaps the polar and azimuthal angles:
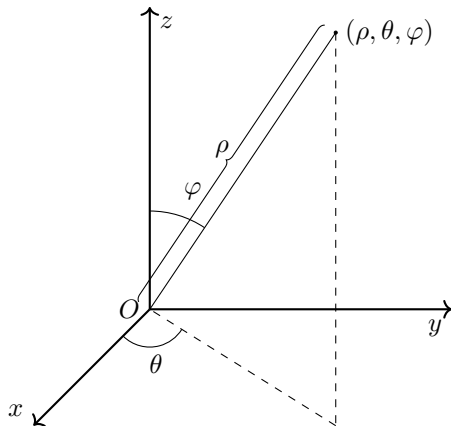
mathematical convention   physics convention

which does not yield a separable solution. For this calculation, it is more convenient to use spherical coordinates. This simplification of the solution comes at the cost of a more complex Schrödinger equation, which is

$$-\frac{\hbar^2}{2m}\nabla^2\psi(r,\theta,\varphi) + U(r)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi)$$

where

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

When the potential energy is only dependent on r, as is the case for the Coulomb potential energy, the solution can be factored as

$$\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$$

where $R(r)$ is the *radial function*, $\Theta(\theta)$ is the *polar function*, and $\Phi(\varphi)$ is the *azimuthal function*. The quantum state of a particle moving in a potential energy that depends only on r can be described by the angular momentum quantum numbers ℓ and m_ℓ.

The polar and azimuthal solutions are given by combinations of standard trigonometric functions while the radial function is obtained by solving the radial equation

$$-\frac{\hbar^2}{2m}\left(\frac{\mathrm{d}^2R}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \left(-\frac{e^2}{4\pi\varepsilon_0 r} + \frac{\ell(\ell+1)\hbar^2}{2mr^2}\right)R(r) = ER(r)$$

Quantum Numbers and Wave Functions

When solving a three-dimensional equation, such as the Schrödigner equation, three parameters emerge as indices for the solutions, just as the single index n applied from the application of boundary conditions to the wave function for the one-dimensional infinite square well. These indices are the three *quantum numbers* that label the solutions. The three quantum numbers that emerge from the solutions and their allowed values are

n	principle quantum number	$1, 2, 3, \ldots$	
ℓ	angular momentum/azimuthal quantum number	$0, 1, 2, \ldots, n-1$	(quantum numbers)
m_ℓ	magnetic quantum number	$0, \pm1, \pm2, \ldots, \pm\ell$	

The principle quantum number is identical to the quantum number n obtained from the Bohr model. It determines the quantized energy levels as

$$E_n = -\frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2}\frac{1}{n^2}$$

Note that the energy only depends on n. The permitted values of ℓ are limited by n and those of m_ℓ are in turn restricted by ℓ.

Complete with quantum numbers, the wave function can be written as

$$\psi_{n,\ell,m_\ell}(r,\theta,\varphi) = R_{n,\ell}(r)\Theta_{\ell,m_\ell}(\theta)\Phi_{m_\ell}(\varphi)$$

The indices are the three quantum numbers necessary to describe the solutions. Some sample wave functions are

n	ℓ	m_ℓ	$R(r)$	$\Theta(\theta)$	$\Phi(\varphi)$
1	0	0	$\dfrac{2}{a_0^{3/2}}e^{-r/a_0}$	$\dfrac{1}{\sqrt{2}}$	$\dfrac{1}{\sqrt{2\pi}}$
2	0	0	$\dfrac{1}{(2a_0)^{3/2}}\left(2-\dfrac{r}{a_0}\right)e^{-r/2a_0}$	$\dfrac{1}{\sqrt{2}}$	$\dfrac{1}{\sqrt{2\pi}}$
	1	0	$\dfrac{1}{\sqrt{3}(2a_0)^{3/2}}\dfrac{r}{a_0}e^{-r/2a_0}$	$\sqrt{\dfrac{3}{2}}\cos\theta$	
		± 1		$\mp\dfrac{\sqrt{3}}{2}\sin\theta$	$\dfrac{1}{\sqrt{2\pi}}e^{\pm i\varphi}$
3	0	0	$\dfrac{2}{(3a_0)^{3/2}}\left(1-\dfrac{2r}{3a_0}+\dfrac{2r^2}{27a_0^2}\right)e^{-r/3a_0}$	$\dfrac{1}{\sqrt{2}}$	$\dfrac{1}{\sqrt{2\pi}}$
	1	0	$\dfrac{8}{9\sqrt{2}(3a_0)^{3/2}}\left(\dfrac{r}{a_0}-\dfrac{r^2}{6a_0^2}\right)e^{-r/3a_0}$	$\sqrt{\dfrac{3}{2}}\cos\theta$	
		± 1		$\mp\dfrac{\sqrt{3}}{2}\sin\theta$	$\dfrac{1}{\sqrt{2\pi}}e^{\pm i\varphi}$
	2	0		$\sqrt{\dfrac{5}{8}}(3\cos^2\theta-1)$	$\dfrac{1}{\sqrt{2\pi}}$
		± 1	$\dfrac{4}{27\sqrt{10}(3a_0)^{3/2}}\dfrac{r^2}{a_0^2}e^{-r/3a_0}$	$\mp\sqrt{\dfrac{15}{4}}\sin\theta\cos\theta$	$\dfrac{1}{\sqrt{2\pi}}e^{\pm i\varphi}$
		± 2		$\dfrac{\sqrt{15}}{4}\sin^2\theta$	

For the ground state ($n = 1$), only $\ell = 0$ and $m_\ell = 0$ are allowed, meaning that the set of quantum numbers for the ground state is $(1,0,0)$. The first excited state ($n = 2$) has 4 possible sets of quantum numbers: $(2,0,0)$, $(2,1,0)$ and $(2,1,\pm1)$. All of their corresponding wave functions describe the same energy, so the $n = 2$ level is *degenerate*. The $n = 3$ level is also degenerate,

having 9 possible sets of quantum numbers. In general, the level with principal quantum number n has degeneracy n^2.

There are several reasons for making the distinction between wave functions that correspond to the same energy. Firstly , the levels are not precisely degenerate, being separated by a small energy of about $10 \times 10^{-5}\,\text{eV}$. Secondly, in studying the transitions between levels, it has been found that the intensities of individual transitions is dependent on the quantum numbers of the particular level from which the transition originates. Thirdly, and perhaps of most import, *each set of quantum numbers corresponds to a very different wave function, therefore representing a very different state of motion of the electron.* These states have different spatial probability distributions, thusly affecting many atomic properties (take, for instance, the way two atoms can form molecular bonds).

Probability Densities

The probability of finding the electron in any spatial interval is determined by the squared amplitude of its wave function. For the hydrogen atom, $|\psi(r,\theta,\varphi)|^2$ gives the *volume probability density* (probability per unit volume) at (r,θ,φ). To compute the actual probability of finding the electron, this must be multiplied by some volume element dV located at (r,θ,φ). In spherical coordinates, this volume element is

$$dV = r^2 \sin\theta \, dr \, d\theta \, d\varphi$$

making the probability of finding the electron in the volume element

$$|\psi_{n,\ell,m_\ell}(r,\theta,\varphi)|^2 = |R_{n,\ell}(r)|^2 |\Theta_{\ell,m_\ell}(\theta)|^2 |\Phi_{m_\ell}(\varphi)|^2 r^2 \sin\theta \, dr \, d\theta \, d\varphi \quad \text{(complete probability density)}$$

7.4   Radial Probability Densities

Rather than considering the probability density to locate the electron, it is advantageous to only consider the probability of finding it at a particular distance from the nucleus, regardless of θ and φ. The *radial probability density* $P(r)$ can be regarded as the probability of finding the electron within a spherical shell of inner radius r and thickness dr centered at the nucleus. This radial probability can be derived from the complete probability by integrating over the domains of θ and φ. This effectively sums the probabilities for the volume elements at a given r for all θ and φ:

$$P(r)\,dr = |R_{n,\ell}|^2 r^2 \, dr \int_0^\pi |\Theta_{\ell,m_\ell}(\theta)|^2 \sin\theta \, d\theta \int_0^{2\pi} |\Phi_{m_\ell}(\varphi)|^2 \, d\varphi$$

These two integrals are simply equal to 1, due to each function being individually normalized, so the radial probability density is simply

$$P(r) = r^2 |R_{n,\ell}(r)|^2 \qquad \text{(radial probability density)}$$

Note that due to the r^2 factor, $P(r)$ must always be 0 at $r = 0$; that is, the probability of locating the electron in a spherical shell goes to 0 as the shell shrinks to radius 0, but the probability density may be nonzero. Moreover, $P(r)$ and $R(r)$ give insight into different aspects of the electron's behavior. For the ground state, for example, the radial wave function is maximized at $r = 0$, but the radial probability density is maximized at $r = a_0$.

Using radial probability densities, the average radial coordinate can be found; that is, the average

distance between the proton and the electron. This is simply the 50th percentile of the probability density; that is,

$$\int_0^{r_{\text{avg}}} P(r)\,\mathrm{d}r = 0.5 \qquad\qquad \text{(average radius)}$$

As in the Bohr model, this average varies roughly with n^2, meaning that an $n = 2$ electron is on average about 4 times further from the nucleus than a ground state electron.

The most probable radius is the location that maximizes $P(r)$. For each n, $P(r)$ for $\ell = n - 1$ has only a single maximum, which occurs at the location of the Bohr orbit:

$$r = n^2 a_0 \qquad\qquad \text{(most probable radius)}$$

7.5   Angular Probability Densities

The angular part of the probability density is obtained by considering only the angular parts of the wave function:

$$P(\theta, \varphi) = |\Theta_{\ell,m_\ell}\Phi_{m_\ell}(\varphi)|^2 \qquad\qquad \text{(angular probability density)}$$

Note that all probability densities are *cylindrically symmetric*, being independent of φ. The wave function for $\ell = 0$ is also *spherically symmetric*, being independent of direction in general.

7.6   Intrinsic Spin

Spatial quantization can be observed by placing an atom in an externally applied magnetic field. The interaction that occurs between the field and the atom's *magnetic dipole moment* (which is related to the electron's angular momentum) enables the observation of the separate components of $\vec{L}$ in addition to ℓ (determined by counting the number of z components, which is equal to $2\ell + 1$). This experiment also yields the electron's unexpected property of *intrinsic spin.*

Orbital Magnetic Dipole Moments

The classical magnetic dipole moment $\vec{\mu}$ is defined as the vector of magnitude equal to the product of the circulating current and the area enclosed by the orbital loop. Its direction is determined by the right-hand curl rule, the fingers curling in the direction of the conventional (positive) current and the thumb indicating the direction of $\vec{\mu}$.

Quantum mechanics forbids precise knowledge of the direction of $\vec{L}$ and therefore that of $\vec{\mu}$ as well. Only their z components can be specified. As the electron is negatively charged, these z components are in opposite directions.

The Bohr model used with a circular orbit yields the relationship between $\vec{L}$ and $\vec{\mu}$, which is identical to the quantum mechanical result. The circulating electron is regarded as a circular loop of current

$$i = \frac{\mathrm{d}q}{\mathrm{d}t} = -\frac{e}{T}$$

where T is the period of the orbit. If the electron is moving with speed

$$v = \frac{p}{m}$$

around a loop of radius r, then

$$T = \frac{2\pi r}{v} = \frac{2\pi r m}{p}$$

The magnitude of the magnetic dipole moment is

$$\mu = iA = -\frac{e}{2\pi r m/p}\pi r^2 = -\frac{e}{2m}rp = -\frac{e}{2m}L$$

Writing this in terms of vectors yields

$$\vec{\mu}_L = -\frac{e}{2m}\vec{L}$$

The negative sign indicates that $\vec{L}$ and $\vec{\mu}_L$ point in opposite directions. The subscript L on the dipole moment indicates that it is a result of the *orbital* angular momentum $\vec{L}$.

The z component of the magnetic moment is

$$\mu_{L,z} = -\frac{e}{2m}L_z = -\frac{e}{2m}m_\ell\hbar = -\frac{e\hbar}{2m}m_\ell = -m_\ell\mu_B$$

where μ_B is the *Bohr magneton*

$$\mu_B = \frac{e\hbar}{2m} \approx 9.27 \times 10^{-24}\ \frac{\text{J}}{\text{T}} \qquad\qquad \text{(Bohr magneton)}$$

Atomic magnetic moments are typically on the order of μ_B, making it a convenient unit for expressing them.

7.7 Energy Levels and Spectroscopic Notation

The introduction of intrinsic angular momentum, or *spin*, requires that a fourth quantum number be introduced. The spin s need not be specified, as it is always $1/2$, being a fundamental property of the electron, but its quantum number m_2 must be specified as $\pm 1/2$, which identifies the z component of the spin. The complete description for an electron's state thus requires the four quantum numbers (n, ℓ, m_ℓ, m_s).

The ground state of hydrogen is now $(1, 0, 0, \pm 1/2)$, making its degeneracy 2. The first excited state now has degeneracy 8. In general, a level's degeneracy is $2n^2$.

The directions (z component) of the angular momentum vectors is important to know when the atom is subject to a magnetic field, but otherwise is of little significance. As such, *spectroscopic notation* is used to label the levels. This system uses letters for the different values of ℓ:

ℓ	0	1	2	3	4	5	6	$\cdots$
Designation	s	p	d	f	g	h	i	$\cdots$

The first four letters stand for sharp, principal, diffuse, and fundamental respectively, which were terms used to describe atomic spectra prior to the development of this theory. Using this notation, the ground state of hydrogen is $1s$, where $n = 1$ is specified before the $\ell = 1$.

Solving the Schrödinger equation and using the solutions to compute *transition probabilities*, it can be found that the most likely transitions to occur are those that change ℓ by 1. This restriction is a *selection rule*:

$$\Delta\ell = \pm 1 \qquad\qquad \text{(selection rule)}$$

Chapter 8

Many-Electron Atoms

In analyzing complex physical systems, it is often convenient to separate the important parts from the less important parts. This approach unfortunately cannot be used in analyzing the motion of electrons in atoms with more than one electron. The problem of mutual significant interactions between 3 or more bodies is called a *many-body problem*. For such problems, exact, closed solutions cannot be found for the Schrödinger equation.

8.1 The Pauli Exclusion Principle

It would be intuitive for all Z electrons to eventually cascade down to the lowest energy level, the $1s$ state. If this were the case, it would be expected that atoms with $Z \pm 1$ display similar properties. Indeed, some properties of atoms, such as the energies of emitted X rays, do vary smoothly with Z. Other properties violate this notion, though, such as reactivity.

The *Pauli exclusion principle* states that *no two electrons in a single atom can share a set of quantum numbers*. The filling of the orbitals in this way accurately models the properties of the elements.

8.2 Electronic States in Many-Electron Atoms

The set of orbits of a specific principle quantum number are the same average distance from the nucleus and is called an atomic *shell*. The shells are designated by letters:

$$\begin{array}{cccccc} n & 1 & 2 & 3 & 4 & 5 \\ \text{Shell} & K & L & M & N & O \end{array}$$

The levels with a specified value for both n and ℓ are *subshells*. The Pauli principle dictates that the maximum allowed electrons in a given subshell is $2(2\ell + 1)$. This is because m_ℓ can take on $2\ell + 1$ values and m_s can take 2 values.

The Periodic Table

The periodic table is an arrangement of the elements in order of ascending Z. They are arranged in such a way that the columns, called *groups*, contain elements with similar properties.

The rules that are followed in the filling of electronic subshells are as follows:

1. The capacity of each subshell is $2(2\ell + 1)$.

2. Electrons occupy the lowest available energy state.

An element's electron configuration is indicated by the subshells in fill order, each of which is denoted by the shell's number followed by the orbital's letter with the superscript of the number of electrons in that shell.
Each row or *period* of the periodic table begins with the next *s* orbital.
The rightmost row of the periodic table is comprised of *noble* or *inert* gases. An element's electron configuration can be denoted by the symbol of the previous noble gas in brackets followed by the subsequent subshells.

Chapter 12

Nuclear Structure and Radioactivity

At an atom's center lies its nucleus, which occupies a mere 10^{-15} of its volume but providing the electric force that holds the atom together. Within the nucleus are positive Z charges. To prevent these charges from repelling each other to the point that they fly apart, the nuclear force must apply an attractive force. This force is the strongest of any known force, providing nuclear binding energies millions of times stronger than those of atoms.

Atomic and nuclear structure share many similarities, but there are two key differences. In atomic physics, the electron is acted upon by an external force, but no such force exists in nuclear physics. The mutual interactions of the nuclear components is what yields the nuclear force, meaning that this many-body problem cannot be simplified. The second difference is that the nuclear force, unlike the Coulomb force, cannot be concisely expressed .

12.1   Nuclear Constituents

The nucleus of an atom of atomic number Z has charge Ze and constitutes 99.9% of the atom's mass. It is also observed that the masses of atoms are, within to about 0.1%, integers. The integer A that describes the number of atomic mass units of an atom is called its *mass number*. It is therefore reasonable to conclude that nuclei are comprised of more fundamental masses each of mass $1\,\mathrm{u}$.

It was (incorrectly) postulated that the nucleus of an atom of mass number A contains A protons. Such a nucleus would have charge Ae rather than Ze, which overshoots the charges of all atoms apart from hydrogen. This difficulty was resolved by the *proton-electron* model, which postulated (again incorrectly) that the nucleus also contained $A - Z$ electrons. Under this model, the mass of the atom would be about A times the mass of the proton, as the masses of the electrons are negligible. This model is not consistent with the uncertainty principle.

Measuring the *very* small effect that the nuclear magnetic moment has on atomic transitions (called the *hyperfine splitting*) yields the result that the proton has an *intrinsic spin* of 1/2. Consider the deuterium nucleus, which under this model contains two protons and one electron, each of which have spin 1/2. The rules for angular momenta dictate that the total spin is then either 1/2 or 3/2. The observed total spin, though, is 1.

This problem was resolved by the discovery of the neutron, a particle of about the same mass as the proton (about 0.1% heavier) but without electric charge. According to the *proton-neutron* model, a nucleus is comprised of Z protons and $A - Z$ neutrons, which yields a nuclear charge of Ze and a mass of A times the mass of the proton (as the proton and neutron have approximately the same mass).

Apart from their electric charges, the proton and neutron are very similar, so they are given the

shared classification of *nucleons*:

Name	Symbol	Charge	Mass	Rest Energy	Spin
Proton	p	$+e$	$1.007\,276\,\text{u}$	$938.28\,\text{MeV}$	$1/2$
Neutron	n	0	$1.008\,665\,\text{u}$	$939.57\,\text{MeV}$	$1/2$

An element's chemical properties are determined by its atomic number Z but not on its mass number A. Different nuclei can have the same Z but different values of A, having the same number of protons but different numbers of neutrons. Atoms of said nuclei have identical chemical properties and are called *isotopes*. An isotope is denoted by

$$^A_Z X_N$$

where X is the symbol of the element, Z is the atomic number, A is the mass number, and $N = A - Z$ is the *neutron number*. The chemical number and symbol both provide the same information, so only one is necessary. A and N also both provide the same information. Often, only X and A are given.

12.2 Nuclear Sizes and Shapes

Like atoms, nuclei lack an easily definable surface or radius. General features of the nuclear density are determined experimentally. As the nuclear force is the strongest force, it may be expected that protons and neutrons are concentrated at the center of the nucleus, but this is not necessarily the case. Instead, the density remains mostly uniform. This gives insight regarding the short range of the nuclear force.

Interestingly, the density of a nucleus does not depend on A. The nucleons per unit volume is approximately constant over the range of nuclei:

$$\frac{\text{number of nucleons}}{\text{volume of nucleus}} = \frac{A}{4\pi R^3/3} \cong \text{constant}$$

assuming the nucleus to be a sphere of radius R. This means that $A \propto R^3$ or $R \propto A^{1/3}$. Defining a constant of proportionality R_0,

$$R = R_0 A^{1/3}$$

R_0 has been experimentally determined as

$$R_0 \approx 1.2 \times 10^{-15}\,\text{m}$$

though the exact value depends on how the radius is defined, ranging from $1 \times 10^{-15}\,\text{m}$ to $1.5 \times 10^{-15}\,\text{m}$. The length $1 \times 10^{-15}\,\text{m}$ is sometimes referred to as one fermi.

What physicists call *nuclear matter* is incredibly dense. While nuclear matter is not found in bulk on Earth, it is present in some stars, in which the gravitational force forces the merging of protons and electrons into neutrons, making a neutron star.

12.3 Nuclear Masses and Binding Energies

Consider a proton and electron at rest separated by a large distance. The total energy of the system is simply the sum of the rest energies of the two particles. Letting the two particles come together

to form a ground-state hydrogen atom, several photons are emitted with energy totaling $13.6\,\text{eV}$, which is the *binding energy* of the atom. Conservation of energy dictates that

$$m_\text{p}c^2 + m_\text{e}c^2 = m(H)c^2 + 13.6\,\text{eV}$$

which can be rewritten as

$$m_\text{p}c^2 + m_\text{e}c^2 - m(H)c^2 = 13.6\,\text{eV}$$

Nuclear binding energy can be similarly formulated. Consider a deuterium nucleus. The nuclear binding energy is

$$B = m_\text{n}c^2 + m_\text{p}c^2 - m_\text{D}c^2$$

where m_D is the mass of the deuterium nucleus. The masses of the proton and the nucleus can be replaced by their corresponding *atomic* masses of

$$m(^1\text{H})c^2 = m_\text{p}c^2 + m_\text{e}c^2 - 13.6\,\text{eV} \qquad \text{and} \qquad m(^2\text{H})c^2 = m_\text{D}c^2 + m_\text{e}c^2 - 13.6\,\text{eV}$$

yielding

$$B = [m_\text{n} + m(^1H) + m(^2H)]c^2 \approx 2.224\,\text{MeV}$$

Replacing the nuclear mass with atomic mass yields

$$B = [m_n + m(^1H) - m(^2H)]c^2$$

Note that the electrons cancel. More generally,

$$B = [Nm_n + Zm(^1_1H_0) - m(^A_ZX_N)]\left(931.5\,\frac{\text{MeV}}{\text{u}}\right)$$

where the masses are the nuclear masses. The binding energy per nucleon is B/A. The separation of nuclei is *fission* while their combination is *fusion*.

Proton and Neutron Separation Energies

Chapter 13

Nuclear Reactions and Applications

13.1 Types of Nuclear Reactions

In a typical nuclear reaction experiment, a beam of particles of type x is fired incident to a target comprised of nuclei of type X. Following the reaction, particles of type y are observed leaving nuclei of type Y. Symbolically,