Additional addenda for "An introduction to special relativity for radiation and plasma physics"

The book "An introduction to special relativity for radiation and plasma physics" by Greg Tallents [7] introduces special relativity with an emphasis on the ways relativity impacts the interaction of light with matter and the production of light by mater. Five additional appendices are presented here: (i) detailing the origin of a common expression used for cosmological redshift, (ii) introducing the effects of gravity on light propagation, (iii) a discussion of electron and photon angular momentum, (iv) an introduction to the Dirac equation, and (v) a brief introduction to the idea behind quantum field theory. These topics are a little beyond the remit of the book, but can be readily understood using the ideas introduced in the book. Errors in the book found or communicated to the author are also listed at the end of this document.

1 The cosmological redshift

The Doppler shift in detected frequencies from distant astrophysical objects is dependent on their velocity of recession from earth. The Doppler effect causes a decrease in the detected frequency of radiation to values smaller than the emitted frequency when the detection frame of reference and the emitting frame of reference are moving apart. Due to the expansion of the universe, radiation from distant parts of the universe is detected on earth after propagating parallel to the velocity between the frames of emission and detection.

In astrophysics and cosmology, the frequencies of emission ω_E and detection ω_{now} are often related to a redshift parameter z defined by

$$1 + z = \frac{\omega_E}{\omega_{now}}.$$

For large cosmic redshift, the redshit $1 + z$ can be shown to be given by the ratio of the cosmic scalelength of the universe now (when the light is detected) and the cosmic scalelength when the light was emitted. In this additional appendix, we show that the cosmological redshift can be derived from the Doppler redshift formulas of Section 2.6 of "An introduction to special relativity for radiation and plasma physics".

In Section 2.6 (Equation 2.48) we found that for a small velocity difference δv along a line-of-sight, the Doppler frequency shift $\delta\omega$ between two frames of reference is given by

$$\frac{\delta\omega}{\omega} = -\frac{\delta v}{c}.$$

In cosmology, velocity differences between frames of reference are determined by the Hubble constant H_0 such that a velocity difference is related to a spatial separation δr by

$$\delta v = H_0\,\delta r.$$

The Hubble constant is the same throughout all space with a current measured value of $H_0 \approx 70$ km s^{-1}/Mpc. The distance unit Mpc here is the megaparsec which is equal to 3.09×10^{19} km.

The Hubble constant is also related to a quantity known as the scalelength a of the universe by

$$H_0 = \frac{da/dt}{a}.$$

The Doppler frequency shift $\delta\omega$ between two frames with velocity difference δv can be related to the change in the scalelength of the universe by

$$\frac{\delta\omega}{\omega} = -\frac{\delta v}{c} = -\frac{H_0\,\delta r}{c} = -H_0\,\delta t = -\frac{(da/dt)\,\delta t}{a} = -\frac{\delta a}{a},$$

where $\delta t = \delta r/c$ is the time for light to travel a distance δr and δa is the change in the scalelength a in the time δt. Moving to infinitessimal increments, we integrate both sides of the relationship $d\omega/\omega = -da/a$. Appropriate limits are ω_E for the frequency of emission and a_E for the scalelength of the universe at the time of emission and ω_{now} for the detected frequency and a_{now} for the scalelength of the universe now. We can write

$$\int_{\omega_E}^{\omega_{now}} \frac{d\omega}{\omega} = -\int_{a_E}^{a_{now}} \frac{da}{a}.$$

Integrating gives

$$\ln\left(\frac{\omega_{now}}{\omega_E}\right) = -\ln\left(\frac{a_{now}}{a_E}\right)$$

which means that

$$\frac{\omega_E}{\omega_{now}} = \frac{a_{now}}{a_E}$$

and that

$$1 + z = \frac{a_{now}}{a_E}.$$

The cosmological result for the redshift parameter z is independent of the intermediate values of the universe scalelength between emission and detection and depends only on the universe scalelength now a_{now} and the universe scalelength a_E at the time of emission. It is sometime implied that as a cosmological redshift may have a gravitational contribution, any derivation requires general relativistic treatments, but this is not needed. For an examination of cosmological spectral shifts in general relativity, see [4].

2 Effects of gravitation on light

General relativity proposes that mass-less photons propagating in vacuum follow trajectories affected by time and space changes caused by the presence of mass. The slowing of time and warping of space are evidenced by the gravitational deflection of light propagation near large masses and by the gravitational frequency shift of light emission near large masses. Forces between masses due to gravity are then explained in general relativity as arising from the warping of space and time.

The frequency shift of light emission near a large mass is known as gravitational redshift. Gravitational redshift occurs when radiation emission near a large mass is shifted to lower frequency when observed well-away from the mass. The decrease in frequency from emission near a large mass when observed away from the mass is a signature of the apparent dilaton of time near a large mass. The observed radiation wavelength from emission near a large mass is increased when observed away from the mass, indicating an apparent increase of spatial length close to the mass.

Our analysis in this appendix makes use of the "equivalence principle" which states that it is not possible to distinguish between an accelerated frame of reference and a frame of reference in a gravitational field. Flight simulators, for example, can accurately emulate the acceleration of an aircraft by tilting the simulator so that trainee pilots fall back in their seats (or lunge forward) just as they would during aircraft acceleration (or de-acceleration). Similarly, the absence of a gravitational field (true weightlessness) is indistinguishable from free-fall acceleration in a gravitational field (apparent weightlessness). "Vomit comet" flights where aircraft descend with acceleration $g = 9.8$ m s^{-2} can achieve weightlessness of the aircraft occupants for several minutes. A person in an enclosed freely falling elevator cannot do an experiment to ascertain if their weightlessness is due to their falling in a gravitiational field or if they are really in outer space away from any gravitational field[1].

It is possible to determine effects of gravity on light using formulas developed for the Doppler effect on light (see Section 2.6) and the propagation directions of light in different frames of reference (see Section 2.5). We consider a falling body in a gravitational field as an inertial frame of reference equivalent to a frame of reference outside of a gravitational field except for a difference in velocity between the two frames. As assumed in Sections 3.3 to 3.6 where we dealt with the relativistic acceleration of bodies, we assume that measurements are made over a sufficiently short period of time that an observer in an accelerating frame of emission moves at close to the same speed as the objects being measured and so remains in the same frame of reference.

Consider a scenario where an observer at a large distance from a large mass records electromagnetic radiation emitted from a body just starting to fall towards the large mass. We assume that the falling body accelerates at a rate g, emits radiation of frequency ω' and that the observer is placed directly above the body so that the radiation propagates opposite to the direction of acceleration. In special relativity, angles of light propagation are measured relative to the velocity $\mathbf{v}$ between frames of reference. The angle θ' of propagation of radiation in the frame of the moving body is π and using Equation 2.43, we see that the angle θ of propagation in the observer frame is also π. Using Equation 2.47 for the Doppler shift, we have that the frequency ω detected is given by

$$\omega = \omega' \gamma \left(1 - \frac{v}{c}\right).$$

A pulse of radiation emitted by the falling body at say time $t = 0$ as it starts to fall is

[1]There are so-called "tidal" forces due to the spherical nature of gravitational fields. Objects at different places in a falling elevator fall towards the center of the mass creating the field rather than exactly parallel to the sides of the elevator. It would be possible, in principle, to measure the distance between objects freely falling with the elevator and if they approach each other, ascertain that the elevator is actually falling in a gravitational field. The elevator would ideally need to be very large and the distance measurement very precise for a useful measurement.

received by the observer at time $t = L/c$ later, where L is the distance between the falling body and the observer. Assuming that in the short period of time for the propagation of the radiation that velocities calculated using Newtonian physics are accurate, the velocity of the observer at the time of detection relative to the falling body is given by $v = g\,t = g\,L/c$. If the velocities v are small, the Lorentz parameter $\gamma \approx 1$ and the detected frequency becomes[2]

$$\omega \approx \omega' \left(1 - \frac{g\,L}{c^2}\right).$$

The radiation detected by an observer outside the gravitational field is shifted to the red (lower frequency).

The accelerating frame of reference of the falling body in the above "thought experiment" is equivalent to a frame of reference in general relativity for a body in a uniform gravitational field. If the "observer" outside the gravitation field were to send an electromagnetic signal to a body in a gravitational field, the signal frequency received by the falling body in the gravitational field is blue-shifted (to higher frequency[3]). The change of the frequency of radiation for the body in the gravitational field relative to the detected frequencies outside the gravitational field can be regarded as occuring due to the apparent dilation of time in the gravitational field. Time t for an observer outside of the field is related to time t' in the field by the reciprocal of the frequency relationship:

$$t \approx \frac{t'}{1 - g\,L/c^2}.$$

This expression is identical to the gravitational time dilation in the limit of a small uniform gravitational field that can be obtained with general relativistic treatments.

In all frames of reference, the standard relationship between frequency ω and wavelength λ is valid:

$$\omega\,\lambda = 2\pi c.$$

Continuing our "thought experiment", the wavelength λ of light for the observer and the wavelength λ' for the falling body in the gravitational field are consequently related by

$$\lambda \approx \frac{\lambda'}{1 - g\,L/c^2}.$$

Wavelengths here are simply measurements of distance parallel to the acceleration g, so we can say that all lengths parallel to the acceleration g in a gravitational field are dilated in this way. This is the same expression for length dilation in the limit of a small uniform gravitational field as obtained with general relativity tensor treatments.

In Section 2.3, we developed expressions for the relativistic velocities of speeding bodies in different frames of reference. Adapting Equation 2.32 to give the transform from the proper frame (primed) to a nominally stationary frame (unprimed), the angle

[2] The light propagates in the frame of the falling body at frequency ω', but is detected in the frame of an observer outside the gravitational field who has an apparent velocity $v = g\,L/c$ away from the frame of the radiation source at the time of detection.

[3] The proper frame of the emitter at frequency ω' is now outside the gravitational field and $\mathbf{v}$ is parallel to the direction of light propagation so that the detected frequency ω in the gravitational field is given by $\omega \approx \omega'(1 + g\,L/c^2)$.

θ of propagation to the velocity $\mathbf{v}$ in a stationary frame is related to the angle θ' to the velocity $\mathbf{v}$ in the moving frame by

$$\tan\theta = \frac{u' \sin\theta'}{\gamma(u' \cos\theta' + v)}.$$

where u' is the speed of the body in the moving frame of reference. A body in a gravitational field with acceleration g can be regarded as being in reference frame that drops with a velocity $v = g\,t$ increasing with time t. In the frame dropping due to gravitational acceleration, the body continues with a velocity u' at an angle θ'. For a velocity perpendicular to the acceleration due to gravity $\theta' = \pi/2$ in the dropping frame and then the angle θ traversed by the body to an observer outside the gravitational field is given by

$$\tan\theta = \frac{u'}{\gamma\, g\, t} = \frac{u}{g\, t}.$$

This expression is, of course, the Newtonian ratio of the horizontal velocity divided by the increasing vertical velocity of the body in the gravitational field. At time $t = 0$, $\theta = \pi/2$ and then θ decreases with time t. In the frame of an observer who is not falling in the gravitational field or an observer outside the gravitational field, the trajectory of the body bends towards the acceleration direction due to gravity. To a falling observer, the body just continues moving with a velocity u'.

We now want to examine the path of an electromagnetic wave in a gravitational field. We first set the propagation direction in the frame of a body dropping in the gravitational field and then determine the propagation direction in a frame outside the gravitational field. We consider light propagating perpendicular to the gravity acceleration in the frame of the body in the field as this has practical implications for light passing close to a large mass. The angle θ of light propagation for a stationary observer is related to the angle θ' in the moving frame by, for example, Equation 2.43 where

$$\cos\theta = \frac{\cos\theta' + v/c}{1 + (v/c)\cos\theta'}.$$

For light propagating at angle $\theta' = \pi/2$,

$$\cos\theta = \frac{v}{c}.$$

To an observer outside the gravitational field, a frame of reference in the gravitational field has a velocity $v = g\,t$, where t represents time. Light propagating perpendicular to the acceleration in a gravitational field in the frame of the falling body, propagates at an angle θ to an observer outside the gravitational field where

$$\cos\theta = \frac{g\,t}{c}.$$

As $c >> g\,t$, θ is close to $\pi/2$ and so we can write that

$$\tan\theta \approx \frac{c}{g\,t}.$$

This is equivalent to the angle θ made by a body with a velocity u moving perpendicular to the gravitational acceleration if we replace u by the speed c of photons. The angle θ is less than $\pi/2$ for $t > 0$ and decreases with time which means that to an observer outside the gravitational field, light propagates at angles less than perpendicular to the acceleration direction and the deviation from perpendicular increases with the time that a photon of light is in the field. Assuming $c >> g\,t$, we can write for the distance y of propagation in the gravitational field that $y \approx c\,t$, so that

$$\tan\theta \approx \frac{c^2}{g\,y}.$$

To an observer outside a gravitational field, the gravitational field "bends" light towards the acceleration direction due to gravity, or in other words, towards a large mass.

As we found for a body with mass traveling slower than the speed of light, the above argument shows that in the frame of an observer outside of a gravitational field, a light trajectory bends towards the large mass. The jump made in the logic of general relativity is to say that the vacuum light path near a large mass defines the deformation of space due to the presence of the mass. This space and time deformation is the same for light and for bodies with mass.[4] The change in frequency of light in a gravitational field to an observer outside the gravitational field represents the apparent slowing of time for all bodies in the field as far as observers outside the field are concerned.

Proximity to a large mass is needed for significant gravitational effect on space and time, but it is possible, for example, to observe small radiation frequency changes at different heights on or near the earth's surface due to the earth's gravity (for the first measurement of earth-bound gravitational redshift, see Pound and Rebka [5]). In 1919, Arthur Eddington measured starlight passing close to the sun during an eclipse and found that the starlight was deflected as predicted by general relativity [2]. The deflection of light by large masses has become a technique to image more distant astronomical objects. A large mass such as a black-hole, dark matter or just a star acts like a lens in collecting and focusing light [1], [3], [8].

The simple treatment presented here can reproduce the form of the deflection formula for light passing a large spherical mass M. Writing that the angle θ of light propagation to the direction of the acceleration g is given by $\pi/2 - \delta\theta$, where $\delta\theta$ is the deflection from perpendicular to the acceleration g, we have that

$$\cos\theta \approx \delta\theta \approx \frac{g\,y}{c^2}.$$

Using the classical Newtonian acceleration g, we have

$$g = \frac{GM}{b^2}$$

where b is the distance from the center of the mass M and G is the gravitational constant. Light propagates for some distance y in the gravitational field around a spherical mass. Taking the distance of propagation in the gravitational field as $y \approx 4b$, where b is now the closest approach to the center of the mass, we obtain that

$$\delta\theta \approx 4\frac{GM}{b\,c^2}$$

[4] Assuming that a mass is sufficiently small that it does not change the gravitational field.

in agreement with general relativity.[5]

3 Electron and photon spin

Fundamental particles such as electrons and photons exhibit an intrinsic angular momentum or spin. Electrons belong to the class of particles with spin $(1/2)\hbar$ known as fermions, while photons are examples of spin $\hbar$ known as bosons. Spin can only be directed with a component parallel to an axis or anti-parallel with values $\pm(1/2)\hbar$ (fermions) or $\pm\hbar$ (bosons). A fundamental definition of the boson class requires the interchange of the position of identical bosons to result in the same spin. Identical bosons in the same state can swap position without changing the spin of the system and so can co-exist in the same state. For fermions, a fundamental definition requires that interchanging the position of two identical fermions leads to a change of spin. However, interchanging identical fermions will result in the same spin, which leads to the conclusion known as the Pauli exclusion principle, that only one fermion at most can occupy a quantum state. The different energy, momentum and velocity distribution functions for fermions (the Fermi-Dirac distribution) and bosons (the Bose-Einstein distribution) resulting from the different possible occupancy of quantum states are discussed in Chapter 6.

In Section 2.9, we found that solutions of the time dependent Schodinger equation for an isolated particle have a temporally varying "phase" of the wavefunction. We can write for the temporal component of the wavefunction

$$\phi(t) = \exp\left(-i\frac{Wt}{\hbar}\right)$$

where W is the energy of the particle. In the frame of the particle, the energy

$$W = m_0 c^2$$

where m_0 is the rest mass of the particle. The temporal component of the wavefunction oscillates in the frame of the particle with a frequency

$$\omega_S = \frac{W}{\hbar} = \frac{m_0 c^2}{\hbar}.$$

Particle probability distributions are determined by the square of the wavefunction amplitude. The temporal oscillation ω_S of the wavefunction of an electron is said to infer a rotation of the electron probability density at a frequency $2\omega_S$ known as zitterbewegung (see Section 2.9).

In order to explore motion of the probability density of an electron, it is reasonable to propose that electrons have a mass distribution over a small volume, rather than having all mass at a "point". We will assume that the electron probability density in the frame of an electron center of mass is symmetric with a Gaussian shape around an axis (such

[5] Our treatment ignores the changing angle of the direction of light propagation to the gravitational acceleration direction for a spherical gravitational field. We define the distance y of light propagation in the field as $y \approx 4b$ to get the correct deflection angle $\delta\theta$. With this assumption, the value of g along the light propagation distance $4b$ varies from the maximum to $1/5^{1/2} \approx 1/2$ of maximum.

as an imposed magnetic field). We assume that the distribution of the electron mass m_0 transverse to the axis can be written as

$$\rho_m(r) = \frac{m_0}{2\pi\,(\Delta r)^2} \exp\left(-\frac{r^2}{2\Delta r^2}\right).$$

Assummg cylindrical symmetry, the factor $m_0/(2\pi\,\Delta r)$ arises as integrating $\rho_m(r)$ over radii from zero to infinity should be normalized to be equal to the particle mass m_0. The standard deviation of an electron position is found by averaging the square of the distance r from the center of mass:

$$\sigma_r^2 = \frac{\int_{r=0}^{\infty} r^2\,\rho_m(r)\,dr}{\int_{r=0}^{\infty} \rho_m(r)\,dr} = \frac{\int_{r=0}^{\infty} r^2\,\exp\left(-\frac{r^2}{2\Delta r^2}\right) dr}{\int_{r=0}^{\infty} \exp\left(-\frac{r^2}{2\Delta r^2}\right) dr} = (\Delta r)^2.$$

The standard deviation σ_r of electron position perpendicular to the axis of symmetry is Δr.

In quantum mechanics, position r and wave number $k = p/\hbar$ are complementary variables, where p is the transverse momentum. An expression for the momentum p transverse to an axis of symmetry is found by taking the Fourier transform of the position distribution $\rho_m(r)$. We have for the momentum distribution

$$\begin{aligned}\rho_m(k) = \rho_m(p/\hbar) &= \frac{1}{m_0}\int_{-\infty}^{\infty} \rho_m(r)\exp(-i\,r\,(p/\hbar))dr \\ &= \frac{2}{2\pi\,(\Delta r)^2}\int_0^{\infty} \exp\left(-\frac{r^2}{2\Delta r^2}\right)\cos\left(\frac{r\,p}{\hbar}\right) dr \\ &= \sqrt{\frac{1}{2\pi}}\left(\frac{1}{\Delta r}\right)\exp\left(-\frac{1}{2}\left(\frac{\Delta r\,p}{\hbar}\right)^2\right). \qquad (1)\end{aligned}$$

The standard deviation Δp of this momentum distribution is given by

$$\left(\frac{\Delta p}{\hbar}\right)^2 = \frac{\int_0^{\infty} k^2\,\rho_m(k)\,dk}{\int_0^{\infty} \rho_m(k)\,dk} = \frac{\int_0^{\infty} k^2\exp(-(\Delta r\,k)^2/2)dk}{\int_0^{\infty} \exp(-(\Delta r\,k)^2/2)dk} = \frac{1}{(2\Delta r)^2}.$$

A Gaussian distribution of momentum can be regarded as arising due to the uncertainty principle with the standard deviation of the linear momentum Δp perpendicular to the axis related to the standard deviation σ_r of the probability density of electron position by

$$\Delta p\ \sigma_r = \Delta p\ \Delta r = \frac{\hbar}{2}.$$

A Gaussian distribution of mass produces this minimum relationship $\Delta p\ \sigma_r$ between the standard deviation of complementary variables. With a Gaussian distribution of momentum, we expect the momentum distribution to vary proportionally to $\exp(-p^2/(2\Delta p)$. Comparing such a distribution to Equation 1, we see that the standard deviation Δp of the transverse momentum is given by the value at radius Δr, that is $\Delta p = 2\omega_S\,m_0\,\Delta r$.

The angular momentum around a point is evaluated by integrating $\mathbf{r} \times \mathbf{p}$, where $\mathbf{r}$ is a vector from the point to a position where $\mathbf{p}$ is the momentum. Ignoring relativistic effects, the velocity of rotation of an electron probability density with radius r from the axis of rotation can be postulated to be given by

$$v(r) = 2\,\omega_S\; r$$

where $2\omega_S$ is the angular rotation velocity which we suppose is such that $\omega_S = m_0c^2/\hbar$. Assuming cylindrical symmetry, the total angular momentum of a rotating mass with distribution $\rho_m(r)$ is

$$\Omega = \int_{r=0}^{\infty} v(r)\, r\, \rho_m(r)\, 2\pi r\, dr = \int_{r=0}^{\infty} 2\omega_S r^2 \frac{m_0}{2\pi\,(\Delta r)^2} \exp\left(-\frac{r^2}{2\Delta r^2}\right) 2\pi r\, dr$$
$$= 2\pi \frac{m_0}{2\pi\,(\Delta r)^2} 2\omega_S \int_{r=0}^{\infty} r^3 \exp\left(-\frac{r^2}{2\Delta r^2}\right) dr = 2\, m_0\, \omega_S\, \Delta r^2$$

We found above that the standard deviation of the transverse momentum Δp is such that

$$\Delta p = 2\omega_S\, m_0 \Delta r.$$

Using this value of Δp and the uncertainty relationship between Δp and σ_r leads to a simplification of the angular momentum of the rotating mass:

$$\Omega = 2\, m_0\, \omega_S\, \Delta r^2 = \Delta p\; \Delta r = \frac{\hbar}{2}.$$

The angular momentum $\hbar/2$ is the amplitude of the intrinsic spin of a fermion. This value of spin is independent of the angular rotation rate $2\omega_S$ and hence is independent of the mass m_0 and the frame of reference (the mass could be "relativistic" γm_0). The spin $\hbar/2$ is also independant of the assumed radius Δr of the particle probability density. Interestingly, setting $\omega_S = m_0c^2/\hbar$ for an electron implies a velocity $v(\Delta r) = c$ at a radius $\Delta r = \lambda_C/(4\pi)$, where $\lambda_C = h/(m_0c)$ is the Compton wavelength (see Exercise 4.13).

There are only two possible directions of spin $(1/2)\hbar$ along an axis produced by, for example, a magnetic field: parallel to the field or anti-parallel to the field. When determining the energy of an electron in the presence of electric or magnetic fields, the spin direction and the above determined spin amplitude can have an effect. For example, much of the fine structure splitting of the energy of quantum states of atoms arises due to the electron spin. Fine structure states typically split in two due to the two possible spin directions when there is a defined axis of symmetry.[6]

Electron spin produces a significant effect when quantum state energies or the path of an electron are affected by a magnetic field. The magnetic moment $\mu_{\mathbf{M}}$ of an orbiting charge is determined by a measure of current flow due to the movement of the charge multiplied by the area enclosed within the current "circuit." We have for a charge q of mass m_0 rotating at speed v at a radius r that the magnetic moment is given by

$$\mu_M = \left(\frac{q\,v}{2\pi\,r}\right) \pi r^2 = \frac{q}{2m_0} m_0\, v\, r = \frac{q}{2m_0} L$$

[6]In hydrogen atoms, quantum states with non-zero electron orbital angular momentum (i.e. orbital quantum number $l > 0$) have an axis of symmetry (the magnetic field produced by the electron orbit) and have fine structure energy splitting due to spin-orbit coupling. States with $l = 0$ have zero orbital angular momentum and no axis of symmetry. The direction and magnitude of the electron spin does not effect the quantum state energies when $l = 0$.

where $L = m_0\, v\, r$ is the angular momentum of the charge. The $q\, v/(2\pi\, r)$ term here is the "current" when the charge q rotates and πr^2 is the area enclosed by the circular current. If the charge q and mass m_0 of a rotating body are uniformly distributed, the magnetic moment of the whole body comprises a summation of all "point" masses and charges and the magnetic momentum $\mu_{\mathbf{M}}$ and angular momentum $\mathbf{L}$ are similarly related by

$$\mu_{\mathbf{M}} = \frac{q}{2\, m_0}\mathbf{L}\,.$$

We can set the magnetic moment and angular momentum to be vectors as they are directed in the same direction for a positive q charge. The energy of a body with a magnetic moment $\mu_{\mathbf{M}}$ in a magnetic field $\mathbf{B}$ is $-\mu_{\mathbf{M}} \cdot \mathbf{B}$, so magnetic moments can be accurately measured when the magnetic field is accurately known. The magnetic moment is usually the most significant measureable effect of the spin of charged particles.

For the electron, measurements show that the amplitude of the magnetic moment parallel to a magnetic field due to spin is given by

$$\mu_M = g_s \frac{e}{2\, m_0}(1/2)\hbar.$$

where g_s is a "spin factor" with a value $g_s = 2.0023$ rather than the value $g_s = 1$ expected assuming the charge of the electron is distributed in the same way as the mass. A more accurate value of $g_s = 2$ close to the experimental value is obtained when the magnetic moment is evaluated using relativistic quantum mechanics (the electron mass and charge can be regarded as spinning at velocities approaching the speed of light). Quantum electrodynamics is needed for a final g_s correction. The g_s value departure from two to 2.0023 arises due to the small probability that an electron can emit and then a short time later re-absorb a virtual photon.

Photons can be regarded as a superposition of states of circular polarization. The rotating electric (and magnetic) fields of circularly polarized light produce an angular momentum or spin. For example, a linearly polarized photon comprises oppositely rotating circularly polarized components of equal amplitude. We can use the diffraction behavior of light to examine the spin of the circularly polarized components of a photon. A diffraction limited beam of light of radius Δd with a Gaussian electric field distribution of form $\exp(-r^2/(\Delta d)^2)$ perpendicular to the direction of propagation diverges in the far-field at an angle θ such that

$$\theta = \frac{\lambda}{\pi\, \Delta d}$$

where λ is the wavelength of the light. The diffraction angle θ is measured from one wing of the beam to the other wing and is determined using the angle where the electric field drops to e^{-1} of the peak electric field. The value of θ can also be related to a component of wave vector Δk directed perpendicular to the propagation direction such that

$$\theta = \frac{2\Delta k}{k}$$

where $k = 2\pi/\lambda$ is the wave vector in the propagation direction. The factor two arises here because θ is the diffracted angle measured from one wing of the beam, through the center of the beam to the other wing. Combining the two expressions for θ, we find that

$$\Delta k = (1/2)k\,\theta = \frac{1}{2}\frac{2\pi}{\lambda}\frac{\lambda}{\pi\,\Delta d} = \frac{1}{\Delta d}$$

The momentum of photons of wave vector $\mathbf{k}$ is given by $\hbar\mathbf{k}$ (see Section 2.8). The angular momentum or spin of a photon can be written as

$$\Omega_p = \Delta d\;\hbar\Delta k = \hbar.$$

The spin $\pm\hbar$ of photons is important when light interacts with atoms. The angular momentum of the system (photon and atom) must be conserved, so that the angular momentum of the atom changes by $\pm\hbar$ when absorbing or emitting a photon.

4 Relativistic quantum mechanics: the Dirac equation

Non-relativistic quantum mechanics using the Schrodinger equation explains the energy levels observed in atoms and ions and is consistent with the behavior of free electrons at moderate velocities. It is possible to examine relativistic effects and effects arising due to electron spin using the Schrodinger equation with relativity and spin treated as perturbations. The energy perturbations due to relativity and spin are small and lead to fine structure quantum states energies with small energy differences from the Schrodinger equation predictions (see Chapter 7 of my book published in 2018 [6]). The Dirac equation introduced in this Section is a more fundamental way to examine atom and ion energies as the Dirac equation incorporates a fully relativistic quantum mechanical treatment of electrons and includes the effects of electron spin.

Ignoring relativistic and particles spin effects, the Schrodinger equation enables calculations of the energies of electrons in bound atoms and ions to a good first approximation. The time dependent Schrodinger equation is of form

$$\hat{H}\Psi = i\hbar\frac{d\Psi}{dt}$$

where Ψ is a wavefunction and $\hat{H}$ is the energy operator or Hamiltonian of form

$$\hat{H} = \frac{\hat{p}^2}{2m_0} + V. \tag{2}$$

Here $\hat{p}$ is a momentum operator and V represents the potential, for example, due to the nuclear charge of an atom or ion. The momentum operator is given by

$$\hat{p} = -i\hbar\nabla\cdot$$

with, for example, components in the (x, y, z) directions of $-i\hbar\,\partial/\partial x$, $-i\hbar\,\partial/\partial y$ and $-\hbar\,\partial/\partial z$. The time independent Schrodinger equation enables a calculation of the energy levels E of an atom or ion using the eigenvalue equation

$$\hat{H}\psi = E\psi$$

where ψ is the time independent wavefunction.

Special relativity requires that energy E and momentum p are related by

$$E^2 = p^2c^2 + (m_0c^2)^2.$$

This special relativity relationship suggests that a Hamiltonian for a particle in a potential V allowing for relativistic effects could take the form

$$\hat{H} = \sqrt{\hat{p}^2c^2 + (m_0c^2)^2} + V. \tag{3}$$

The interpretation of the square root in operator notation is not clear. This led Paul Dirac (1902 - 1984) to propose an operator consistent with relativity of form

$$\hat{H} = \left(c \sum_{j=1}^{3} \alpha_j \hat{p}_j \right) + \alpha_0 \, m_0 c^2 + V \tag{4}$$

where α_j ($j = 0, 1, 2,$ and 3) are to be determined. It transpires that the α_j values here cannot be numbers (complex or real) as squaring the Dirac proposal for $\hat{H}$ (Equation 4) to get back to the function within the square root of Equation 3 requires, for example, that $\alpha_j\alpha_k + \alpha_k\alpha_j = 0$ ($k \neq j$) which is not possible with numbers. Matrices can have this non-commutative relationship so Dirac proposed that the α_j are matrices.

Mathematical treatments show that the α_j must be even column and row matrices (that is 2×2, 4×4, ...) which then requires the wavefunctions to be a vector of two or four row length. The chosen matrices α_j are, in fact, 4×4 and the wavefunction vectors have four rows. It is convenient to write down the appropriate values of α_j as 2×2 matrices, but with each element representing an internal 2×2 matrix. For example

$$\alpha_0 = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

The values of α_j ($j = 1, 2$ and 3) have elements which are known as the Pauli spin matrices and encode the effect of particle spin. We can write that

$$\alpha_j = \begin{bmatrix} \mathbf{0} & \sigma_{\mathbf{j}} \\ \sigma_{\mathbf{j}} & \mathbf{0} \end{bmatrix}$$

where $j = 1, 2, 3$ represents components in (x, y, z) coordinates such that

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$\sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\sigma_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

The Dirac Hamiltonian can be written as

$$\hat{H} = m_0 c^2 \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} + c \sum_{j=1}^{3} \left[\hat{p}_j \begin{bmatrix} \mathbf{0} & \sigma_{\mathbf{j}} \\ \sigma_{\mathbf{j}} & \mathbf{0} \end{bmatrix} \right] + V \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \tag{5}$$

or more briefly as

$$\hat{H} = (m_0 c^2 + V)\alpha_0 + c\,\alpha \cdot \hat{\mathbf{p}} \tag{6}$$

where α and $\hat{\mathbf{p}}$ are respectively vectors $(\alpha_1, \alpha_2, \alpha_3)$ and $(\hat{p}_x, \hat{p}_y, \hat{p}_z)$ corresponding to the (x, y, z) coordinates.[7]

The kinetic energy of a particle of charge q and velocity $\mathbf{v}$ in a magnetic field changes in time with the vector potential $\mathbf{A}$ as $-q\,d(\mathbf{v} \cdot \mathbf{A})/dt$ (see Section 5.1). In an atom or ion, we can incorporate effects of generated magnetic fields for electrons of charge $-e$ by adding an additional term to the Dirac Hamiltonian:

$$\hat{H} = (m_0 c^2 + V)\alpha_0 + c\,\alpha \cdot (\hat{\mathbf{p}} + e\mathbf{A}) \tag{7}$$

where $\mathbf{A}$ is the vector potential with components (A_x, A_y, A_z). Practical solutions of Equation 7 for atoms and ions require conversion of the equation to spherical polar coordinates. A complete relativistic Hamiltonian incorporating the effects of electron spin can then be evaluated. For the hydrogen atom and hydrogen-like ions, we find that

$$\hat{H} = \frac{\hat{p}^2}{2m_0} + V(r) - \frac{\hat{p}^4}{8m_0^3} + \frac{1}{2m_0^2 c^2} \frac{1}{r} \frac{dV(r)}{dr} \hat{S} \cdot \hat{L} + \frac{\hbar^2}{8m_0 c^2} \nabla^2 V(r) \tag{8}$$

with $V(r) = -1/(4\pi\epsilon_0)\, e/r^2$. The first two terms represent the Hamiltonian neglecting relativistic effects and spin (see Equation 2). The remaining terms give energy corrections known as fine structure. The third term $\hat{p}^4/8m_0^3$ is a relativistic correction, while the fourth term is the effect of spin-orbit coupling. Here $\hat{S}$ is a spin operator and $\hat{L}$ an electron orbital angular momentum operator. The energy of the bound electron in hydrogen is changed slightly dependent on the direction of electron spin relative to the orbital angular momentum of the electron. The last term is the Darwin correction to the Hamiltonian. The only remaining untreated effects on hydrogen atom energies are a small quantum electrodynamic effect known as the Lamb shift which effects states with zero orbital angular momentum (s-states) and an even smaller hyperfine effect due to the magnetic moment of the nucleus.

5 A brief introduction to quantum field theory

Quantum field theory (QFT) allows an understanding of the fundamental nature of particles. QFT combines classical field theory, special relativity, and quantum mechanics to enable the construction of physical models of subatomic particles. In “An introduction to special relativity for radiation and plasma physics”, there is an emphasis on the properties of the electron which at the QFT level forms the study of quantum electrodynamics. QFT is a more general treatment for all particles.

[7]The components $(\alpha_1, \alpha_2, \alpha_3)$ of the vector α are 4×4 matrices.

QFT treats particles as excited states (or quanta) of an underlying quantum field. There are different fields for each type of particle. For example, electrons, quarks and photons all have an associated field. We saw (Section 6.4) that electromagnetic radiation is quantised in units of photons in a radiation field with each excitation of a harmonic oscillator expression representing a photon unit. This idea is extended in QFT to particles other than the photon.

Photon propagation is governed by the wave equation (e.g. Equation 6.39). We can propose that the propagation of other particles is also governed by a similar equation. Re-arranging Equation 6.39 to the form of a wave equation gives

$$c^2\,\nabla^2\psi - \frac{\partial^2\psi}{\partial t^2} = 0.$$

We introduce an extra term to this wave equation, such that

$$c^2\,\nabla^2\psi - \frac{\partial^2\psi}{\partial t^2} = \omega_S^2\psi. \tag{9}$$

where the $\omega_S^2\psi$ term introduces a "resonance" to the wave equation oscillations. Solutions of wave equations oscillate in space and time such that

$$\psi(\mathbf{r},t) = A\exp\left(i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right).$$

where $\mathbf{k}$ represents the spatial frequency of oscillation, ω the temporal (angular) frequency of oscillation and A is the amplitude of oscillation. Substituting into Equation 9 yields that

$$\omega^2 = \omega_S^2 + c^2k^2$$

We have that the frequency ω of oscillation is given by

$$\omega = \left(\omega_S^2 + c^2k^2\right)^{1/2}. \tag{10}$$

For photons with $\omega_S = 0$, substituting $k = 2\pi/\lambda$ for the wavelength λ and $\nu = 2\pi\omega$ for the frequency of the oscillating solution yields $\nu\lambda = c$. For $\omega > 0$, we see that

$$\omega \geq \omega_S$$

The minimum frequency of oscillation is ω_S when the extra term is added to Equation 9.

As for the harmonic oscillator solution developed in Section 6.8, the harmonic oscillator energies of the quantum solutions of our wave equation (Equation 9) are quantised in units of $\hbar\omega$. The energies E_T of oscillation are given by

$$E_T = (n + 1/2)\,\hbar\omega$$

where $n \geq 0$ is an integer. The $(1/2)\,\hbar\omega$ value obtained when $n = 0$ represents a zero-point energy, sometimes quoted as the energy of the vacuum. The energy E_n of the harmonic oscillator in excess of the zero-point energy is obtained by multiplying the frequency ω_S and ck term as found above ($\omega^2 = \omega_S^2 + (ck)^2$) by the reduced Planck's constant $\hbar = h/2\pi$. We have

$$E_n^2 = (\hbar ck)^2 + (\hbar\omega_S)^2.$$

The minimum energy E_1 is $\hbar\omega_S$ as the minimum frequency is ω_S. As the harmonic oscillator energies are equally spaced in value, the energies E_n are all given by

$$E_n = n\hbar\omega_S.$$

The energies E_n can also be equated to the energy of a particle in special relativity (see Section 2.8). When p is the momentum of a particle and m_0 is the mass of a particle, we use one of the key results of special relativity to write that

$$E^2 = (p\,c)^2 + (m_0c^2)^2$$

Equating $(\hbar\omega_S)^2$ and $(m_0c^2)^2$ terms in the respective expressions for E_n and E means that

$$\hbar\omega_S = m_0c^2$$

as found in Section 2.9. The value of m_0c^2 is the rest mass energy of a particle of mass m_0. Equating the $\hbar ck$ and pc terms gives the momentum of a particle in the form of the De Broglie relationship:

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

The quantization of the harmonic oscillator solutions shows that the minimum energy (above the zero-point energy) has a value m_0c^2 and increases in units of m_0c^2. In QFT, each quanta of energy in the underlying quantum field for the particle represents an identical particle of rest mass energy m_0c^2. In QFT, the universe is said to be permeated by the underlying quantum field for each particle. Most space has $n = 0$, that is the field exhibits only the vacuum energy $(1/2)m_0c^2$ of the particle. Associated with the uncertainty principle for energy/time, virtual particles exist for short durations of time ($\approx \hbar/m_0c^2$) in the vacuum. In the vacuum, excitation to the $n = 1$ state (creating a virtual particle) is rapidly followed by decay back to $n = 0$.

QFT enables studies of many particles (many field excitations) and the interactions between particles, whereas quantum mechanics using, for example, the Schrodinger equation is usually limited to the evaluation of the energy of a single particle. The idea of a quantum field for each particle can be used to explain the creation and annihilation of particles. For particles with mass, creation/annihilation largely occurs in the early universe, in cosmic ray interactions with the atmosphere and in particle physics experiments. The creation or annihilation of a particle is represented by an excitation or decay between the harmonic oscillator solutions appropriate to the underlying quantum field. Energy conservation requires that new particles are created by decay of particles with a higher sum of rest mass plus kinetic energy.

Interactions between particles are modeled in QFT by the interaction of the underlying quantum fields of the particles. For example, two electrons can interact by one electron emitting a virtual photon that is absorbed by the other electron. The exchange of photon momentum $\hbar\omega/c$ causes the electrons to move apart as we would expect for two similarly charged particles. Electromagnetic forces between particles are said to be mediated by photon interactions. In atomic and particle physics, different bosons of spin 1 mediate different interactions. Photons mediate electromagnetic interactions, the W and Z bosons mediate the weak nuclear interactions, while gluons mediate the strong interaction between quarks. QFT has developed to produce a comprehensive theory of elementary particles and their interactions as evidenced, for example, by the three volumes authored by the 1979 Nobel prize winner Steven Weinberg [9].

6 Errata

Errors in the book ”An introduction to special relativity for radiation and plasma physics“ uncovered or communicated to the author are listed here.

6.1 p42. footnote

The momentum values quoted in the footnote are incorrect and should read that $p_x = \gamma m_0 v$ and $p_y = \gamma m_0 v$.

6.2 p52. Definition of tensor rank

A 4×4 tensor with 16 elements should be referred to as a second rank tensor, while column contravariant and row covariant vectors can be referred to as first rank tensors.

References

[1] S. Birrer and A. Amara. Lenstronomy: Multi-purpose gravitation lens modelling software package. *Phys. of the Dark Universe*, 22:189 – 201, 2018.

[2] F. W. Dyson, A. S. Eddington, and C. Davidson. IX. a determination of the deflection of light by the sun's gravitational field from observation made at the total eclipse of may 29, 1919. *Phil. Trans. R. Soc. Lond.*, 220:291 – 335, 1920.

[3] H. Miyatake et al. First identification of a CMB lensing signal produced by 1.5 million galaxies at $z \approx 4$: constraints on matter density fluctuations at high redshift. *Phys. Rev. Lett.*, 129:061301, 2022.

[4] J. V. Narlikar. Spectral shifts in general relativity. *Am. J. Phys.*, 62:903 – 907, 1994.

[5] R. V. Pound and G. A. Rebka. Gravitational red-shift in nuclear resonance. *Phys. Rev. Lett.*, 3:439 – 441, 1959.

[6] G. J. Tallents. *An introduction to atomic and radiation physics of plasmas.* Cambridge University Press, Cambridge, U.K., 2018.

[7] G. J. Tallents. *An introduction to special relativity for radiation and plasma physics.* Cambridge University Press, Cambridge, U.K., 2022.

[8] D. Walsh, R. F. Carswell, and R. J Weymann. 0957 + 561 A, B: twin quasistellar objects or gravitational lens? *Nature*, 79:381 – 384, 1979.

[9] S. Weinberg. *The quantum theory of fields Vol. 1 - 3.* Cambridge University Press, Cambridge, U.K., 2000.