Notes on relativistic quantum mechanics

Module 4 of Refresher course conducted by Indian Academies of Sciences
St Berchman's college, Changanacherry, Kerala, May 8-14, 2013

Govind S. Krishnaswami, Chennai Mathematical Institute

• These are very brief and incomplete notes based on lectures given at the above Refresher Course. Please let me know (govind@cmi.ac.in) of any comments or corrections. 25 Oct, 2016.

Contents

1 References

• Here are some books where one may read about relativistic quantum mechanics.

1. P J E Peebles, Quantum Mechanics, Princeton Univ Press (1992), chapt 8.
2. B H Bransden and C J Joachain, Quantum Mechanics, 2nd Ed., Pearson (2000), Chapt 15.
3. L I Schiff, Quantum Mechanics, 3rd Ed. McGraw-Hill (1968), Chapt 13.
4. J J Sakurai, Advanced Quantum Mechanics (1967), Chapt 3.
5. R Shankar, Principles of Quantum Mechanics, 2nd Ed. Plenum Press (1994) Chapt 20.
6. P M Mathews and K Venkatesan, A textbook of quantum mechanics, Tata McGraw-Hill (1977) Chapt 10.
7. P A M Dirac, The principles of quantum mechanics, 4th Ed., Oxford (1958), Chapt 11.
8. R Parthasarathy, Relativistic Quantum Mechanics, Narosa (2009).

2 Schrödinger-Pauli equation for spin half electron

• A spin half non-relativistic particle (electron) is described by a wave function $\psi_{\pm}(x)$ which gives the probability amplitude of finding the particle at x with spin projection $S_z = \pm\frac{1}{2}\hbar$.

This wave function lives in the Hilbert space $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. The hamiltonian for such a particle in a potential V is $H = (\frac{p^2}{2m} + V(x)) \otimes I$ and is proportional to the identity in spin space (spin-independent). Now if such a particle has charge e and is in a magnetic field, by analogy with the classical dipole interaction, we add the spin magnetic dipole moment energy

$$H_{md} = -g\frac{e}{2m}\vec{S}\cdot\vec{B} = -\frac{ge\hbar}{4m}\vec{\sigma}\cdot\vec{B} \approx -\frac{e\hbar}{2m}\sigma\cdot B. \tag{1}$$

The dimensionless factor g in the spin gyromagnetic ratio is $g \approx 2$ for the electron, based on experiment. Note that H_{md} is not just proportional to the identity in spin space, though it acts as the identity on the translational degrees of freedom (position-independent but spin dependent)

• On the other hand, we know that the interaction of a charged (spin zero) particle with electromagnetic fields is given by replacing $\vec{p} \to \vec{p} - e\vec{A}$ and $E \to E - e\phi$ in the hamiltonian $H = p^2/2m + V$

$$\left(\frac{p^2}{2m} + V\right)\psi = \hat{E}\psi \longrightarrow \left(\frac{(p-eA)^2}{2m} + V + e\phi\right)\psi = \hat{E}\psi = i\hbar\frac{\partial\psi}{\partial t} \tag{2}$$

Now we wish to generalize this hamiltonian to the case of a spin half particle. The hamiltonian cannot simply be proportional to the identity in spin space as that would not give rise to a magnetic moment interaction, which we expect to arise as a consequence.

• Notice that for a free particle, the hamiltonian $H = \frac{p^2}{2m} \otimes I$ could equally well be written $H = \frac{1}{2m}(\vec{\sigma}\cdot\vec{p})^2$ on account of $\sigma_i\sigma_j = \delta_{ij} + \sqrt{-1}\epsilon_{ijk}\sigma_k$. This suggests a hamiltonian for a spin-half charged particle in an electromagnetic field, acting on two-component spinor wave functions

$$H = \frac{1}{2m}(\sigma\cdot(p-eA))^2 + e\,\phi\otimes I. \tag{3}$$

The corresponding Schrödinger equation for this hamiltonian is called the Pauli equation (1927). Of course, it is just a guess for the appropriate hamiltonian. But it is a good guess. To see why, we use the above identity in the form $(\sigma\cdot A)(\sigma\cdot B) = A\cdot B + i\sigma\cdot(A\times B)$ to write (show this!)

$$H = \frac{1}{2m}(p-eA)^2 + \frac{i}{2m}\sigma\cdot(p-eA)\times(p-eA) + e\phi = \frac{1}{2m}(p-eA)^2 - \frac{e\hbar}{2m}\sigma\cdot B + e\phi. \tag{4}$$

In addition to the usual (spin-independent) electromagnetic interactions we also get the expected spin magnetic moment coupling with the approximately correct gyromagnetic ratio $g = 2$ for the electron.

3 Relativistic quantum mechanics

• The Schrödinger and Schrödinger-Pauli equations *with the above hamiltonians* can be used to describe non-relativistic particles of spin zero (no internal degrees of freedom) and half. Schrodinger in 1926 looked for a wave equation that was appropriate to a particle that might travel at speeds approaching that of light. Though he had an electron in mind, he did not consider its spin as that concept was still being developed. He obtained a relativistic wave equation now called the Klein-Gordon (KG) equation; it is of some relevance to a spin zero

particle like a pion. A relativistic wave equation relevant to a spin half particle was subsequently discovered by Dirac (1928).

• In hindsight, both the KG and Dirac relativistic wave equations are flawed in that they cannot provide a self-consistent description of a single relativistic particle. The number of particles is not a conserved quantity in a relativistic setting, due to the processes of particle production and annihilation. Despite their inconsistencies, these equations are approximately valid when appropriately interpreted in the context of many-particle quantum mechanics. They also lead to many correct physical predictions, such as the existence of anti-particles and relativistic corrections to the hydrogen spectrum and approximately correct electron gyromagnetic ratio.

• On the other hand, the Schrödinger eqn $i\hbar\frac{\partial\psi}{\partial t} = H\psi$ has stood the test of time. For appropriate hamiltonians and Hilbert spaces, it is, to this day, believed to exactly describe both relativistic and non-relativistic situations in the framework of quantum field theory. But it requires more sophistication (and is rather cumbersome) to derive the strikingly simple physical results of the KG and Dirac equations, by solving the SE for a quantum field. So, for historical and practical reasons, we discuss the KG and Dirac equations in attempts to develop a relativistic quantum mechanics of spin-0 and $\frac{1}{2}$ particles.

3.1 Klein-Gordon equation

• Recall that the Schrödinger equation could be obtained by the so-called correspondence rule. This involves replacing $E \to i\hbar\frac{\partial}{\partial t}$, $\vec{p} \to -i\hbar\nabla$ and $x \to \hat{x}$ in the expression $H = E$ where H is the classical hamiltonian of a non-relativistic particle. Then we ask that the quantum wave function satisfy the resulting differential equation:

$$E = \frac{p^2}{2m} + V(x) \quad \longrightarrow \quad i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(x)\psi(x). \tag{5}$$

We want to do the same thing for a free massive relativistic particle, whose energy is given by $E = \sqrt{p^2c^2 + m^2c^4}$. The resulting differential equation is

$$i\hbar\frac{\partial\psi}{\partial t} = \sqrt{-\hbar^2c^2\nabla^2 + m^2c^4}\ \psi \tag{6}$$

To make sense of the operator on the rhs, we could expand in inverse powers of m (around the non-relativistic limit)

$$i\hbar\frac{\partial\psi}{\partial t} = mc^2\left[1 - \frac{\hbar^2\nabla^2}{2m^2c^2} - \frac{\hbar^4\nabla^4}{8m^4c^4} + \ldots\right]\psi. \tag{7}$$

At leading order we get the non-relativistic SE where the hamiltonian includes an additive constant mc^2 coming from the rest energy. However, this new equation is first order in time derivatives but of infinite order in space derivatives. It is not a differential equation in the usual sense, making it a bit hard to work with. Moreover, relativistic covariance is not manifest since space and time derivatives appear very differently. The above equation could be useful in working out relativistic corrections to the SE by truncating the series. But due to the difficulties in dealing with a differential operator of infinite order, and lack of manifest relativistic covariance, a simpler relativistic wave equation was investigated.

• The relativistic energy-momentum relation can also be written $E^2 = p^2c^2 + m^2c^4$, though this includes negative energies $E = \pm\sqrt{p^2c^2 + m^2c^4}$, which are not admissible in classical physics.

But we set aside this objection for now and work with this form due to its simplicity, largely on aesthetic grounds. Applying the correspondence rule, we get the Schrödinger relativistic wave equation or (massive) Klein-Gordon equation

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = -\hbar^2c^2\nabla^2\psi + m^2c^4\psi \quad \text{or} \quad \left(\hbar^2\Box + m^2c^2\right)\psi = 0. \tag{8}$$

$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ is the d'Alembert or wave operator. Notice that the KG equation is 2^{nd} order in both space and time derivatives unlike the SE which is first order in time. Notice that the KG equation admits every solution of (6) as a solution. So even if our aim was to exclusively study (6), it could be technically easier to do so by solving the KG equation first and then discarding the 'unwanted' solutions. But we wish to study the KG equation in its own right for now, even if only for *aesthetic* reasons.

• To decide whether KG is a physically correct equation and what it might physically describe, we need to study its features. With the benefit of hindsight, we can say that suitably interpreted, the predictions of KG are in reasonable agreement with experimental findings concerning relativistic spin-less particles (such as pionic atoms). Ultimately, this is the *physical* justification to study it. The KG equation also forms the foundation of scalar quantum field theory upon second quantisation. In that form and with some additional features, it is used to describe the Higgs field in particle physics.

3.1.1 Plane wave solutions of KG

• To get a feeling for what the KG equation describes, let us look for separable solutions $\Psi(r,t) = \psi(\vec{r})T(t)$. We find

$$-\hbar^2\frac{\ddot{T}}{T} = \frac{m^2c^4\psi - \hbar^2c^2\nabla^2\psi}{\psi} = E^2 \tag{9}$$

where we introduced a separation constant E^2 independent of both $\vec{r}$ and t and having dimensions of energy-squared. $\psi(\vec{r})$ must be an eigenfunction of the Laplacian (i.e., satisfy the Helmholtz equation)

$$\left(-\hbar^2c^2\nabla^2 + m^2c^4\right)\psi = E^2\psi. \tag{10}$$

The operator $-\hbar^2c^2\nabla^2 + m^2c^4$ coincides with the hamiltonian of a non-relativistic particle in a constant potential. So its eigenfunctions are free particle energy eigenstates (e.g. plane waves propagating in any direction). It is a positive operator. So the separation constant E^2 must be positive, which justifies the notation E^2 with E real. This then guarantees that the solutions be oscillatory in time, for if we denote by E either the positive or negative square-root of E^2, then we have

$$T(t) = Ae^{iEt/\hbar} + Be^{-iEt/\hbar} \tag{11}$$

Now $E^2 - m^2c^4$, being the eigenvalue of a positive operator $-\hbar^2c^2\nabla^2$ must be non-negative. Let us denote the quantity $E^2 - m^2c^4$ by p^2c^2 for some positive number p^2. We of course recognize that the above Helmholtz equation arises from the relativistic energy-momentum dispersion relation $E^2 - m^2c^4 = \vec{p}^2c^2$ upon use of the correspondence rule $\vec{p} \to -i\hbar\nabla$. The general solution of the Helmholtz equation is a linear combination

$$\psi(\vec{r}) = Fe^{i\vec{p}\cdot\vec{r}/\hbar} + Ge^{-i\vec{p}\cdot\vec{r}/\hbar} \tag{12}$$

where $\vec{p}$ is *any* 'momentum' vector that satisfies the so-called mass-shell condition $c^2\vec{p}^2 = E^2 - m^2c^4$. The mass shell condition defines a hyperboloid in the space of 4-momenta.

Thus, separable solutions of the KG equation take the form

$$\Psi(\vec{r},t) = \left(Fe^{i\vec{p}\cdot\vec{r}/\hbar} + Ge^{-i\vec{p}\cdot\vec{r}/\hbar}\right)\left(Ae^{iEt/\hbar} + Be^{-iEt/\hbar}\right) \tag{13}$$

These solutions are bounded over all of space at all times, and to that extent, could potentially describe the amplitude of some disturbance. We may synthesize the general solution of the KG equation by taking an arbitrary linear combination of separable solutions. But a peculiar feature is that for a fixed momentum vector $\vec{p}$, there is both a plane wave that moves in the direction of $\vec{p}$ and in the opposite direction. This is at variance with our experience from classical mechanics as well as non-relativistic quantum mechanics where a particle with given momentum moves in the direction of the momentum vector. The current peculiarity is a reflection of the fact that we started with $E^2 - \vec{p}^2c^2 = m^2c^4$, which includes both positive and negative energies for a given momentum vector. This problem did not arise for the Schrödinger equation as it is first order in time, while the KG equation is second order in time.

• Another way of looking at this: The KG equation admits (check by substitution) plane wave solutions $e^{i(\vec{k}\cdot\vec{r}-Et/\hbar)}$ where $\vec{k} = \vec{p}/\hbar$ is an arbitrary wave vector and

$$E = \pm\sqrt{m^2c^4 + \hbar^2c^2\vec{k}^2}. \tag{14}$$

E may be called energy. We may call the mode with $E > 0$ or $\omega = E/\hbar > 0$ a positive energy/frequency mode and one with $\omega < 0$ a negative energy mode. This nomenclature is somewhat arbitrary since we could have written the plane wave as $e^{i(\vec{k}\cdot\vec{r}+Et/\hbar)}$. Independent of convention, for every plane wave with 'energy' E, there is one with energy $-E$. In this sense, the spectrum of energies of the massive KG equation is continuous and comes in two disjoint pieces $(-\infty, -mc^2] \cup [mc^2, \infty)$. So the energy spectrum is not bounded below, there is no least value of E.

• Calculate the group speed $c_g = \frac{\partial\omega}{\partial k}$ of disturbances that propagate according to the dispersion relation $E = \hbar\omega(k) = \sqrt{m^2c^4 + \hbar^2k^2c^2}$. Show that the group speed is less then the speed of light and approaches c when $m \to 0$. The group speed is the speed at which signals propagate. The phase speed $c_p = \omega/k$ can exceed the speed of light. Physical signals do not travel at the phase speed. The different planes waves that combine to form a wave packet can have phase speeds that exceed the speed of light, they destructively interfere at most locations except in the vicinity of the wave packet, which travels at the group speed.

• One option is to simply disallow the negative energy solutions. For example, we might implement this for plane waves by allowing only those initial conditions which ensure that the wave moves in the direction of the momentum vector, ensuring that $E > 0$. Within the context of the KG equation, this is seemingly ok, since the particle will then remain in that stationary state for ever. However, under the influence of external perturbations, the particle could make a transition to a lower energy state. And since there is no ground state, the particle could keep dropping down in energy while emitting radiation. The system is unstable to perturbations as it does not have a ground state. This is problematic since we could extract an infinite amount of energy from such a particle as it makes transitions to states of arbitrarily negative energy.

• Despite this difficulty with trying to interpret solutions of the KG equation as the wave

function of a particle, the equation exhibits several physically desirable features, such as a local conservation law and Lorentz invariance, which we describe next.

3.1.2 Lorentz invariance of the Klein-Gordon equation

The principles of special relativity say that there is no way of physically distinguishing between different frames of reference related by Lorentz transformations. A way of ensuring this is for the differential equations describing the laws of physics to take the same form in all such frames i.e., to be Lorentz invariant/covariant. For example, the terms in an equation could all be scalars (no-uncontracted indices), in which case the equation would be Lorentz invariant (KG is an example as we will see below). More generally the terms could all be tensors of the same sort (e.g. one uncontracted upper index - a vector equation as in the case of Maxwell's equations $\partial^\mu F_{\mu\nu} = j_\nu$), in this case the equation is said to be Lorentz covariant and to transform as a vector. Negative energy solutions are in a sense the price we have to pay for manifest Lorentz covariance.

• To discuss the Lorentz invariance of KG, we define the Minkowski metric $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ and the 4-vector coordinate and gradient

$$x^\mu = (x^0, \vec{x}) = (ct, \vec{x}), \quad \partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right). \tag{15}$$

A Lorentz transformation[1] $x' = \Lambda x$ is one that preserves inner products $\langle x, y\rangle = x^t \eta y$ of 4-vectors: $\langle x, y\rangle = \langle x', y'\rangle$. So Lorentz transformations preserve lengths and angles between 4-vectors. This is the condition that for any 4-vectors x, y

$$x^t \Lambda^t \eta \Lambda y = x^t \eta y. \tag{16}$$

In other words, the Lorentz transformation matrix must satisfy $\Lambda^t \eta \Lambda = \eta$. In components, this reads

$$x'^\mu = \Lambda^\mu_\nu x^\nu \quad \text{and} \quad \Lambda^\mu_\nu \, \eta_{\mu\rho} \, \Lambda^\rho_\sigma = \eta_{\nu\sigma}. \tag{17}$$

We say that the Lorentz transformation preserves the metric.

• x^μ is called a contravariant vector or the contravariant components of the position and ∂_μ is a covariant vector or covariant components of the gradient. The terminology is because of the way they behave under a Lorentz transformation (transforming via Λ and Λ^t):

$$x'^\mu = \Lambda^\mu_\nu x^\nu \quad \text{and} \quad \partial'_\mu = \Lambda^\nu_\mu \partial_\nu \tag{18}$$

Indices are raised by the (inverse) Minkowski metric $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$ and lowered by its inverse $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. The vectors with raised/lowered indices are denoted by the same symbols

$$x_\mu = \eta_{\mu\nu} x^\nu = (ct, -\vec{x}), \quad \partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right) \tag{19}$$

[1] For example, a Lorentz boost in the x direction leads to $x' = \gamma(x - vt), t' = \gamma(t - xv/c^2), y' = y, z' = z$ and $E' = \gamma(E - p_x v), p'_x = \gamma(p_x - vE/c^2), p'_y = p_y, p'_z = p_z$. The nontrivial part of the corresponding transformation, in matrix form, is $\begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$ and $\begin{pmatrix} E'/c \\ p' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} E/c \\ p \end{pmatrix}$ so that $\Lambda = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix}$ where $\gamma = (1 - \beta^2)^{-1/2}$ and $\beta = v/c$. A Lorentz transformation matrix is symmetric for a boost but not for a rotation.

are the covariant components of the position 4-vector and the contravariant components of the gradient.

• Then it is seen that the d'Alembert wave operator is $\Box = \partial_\mu \partial^\mu$. The wave operator is Lorentz invariant as it is the inner product of a 4-vector with itself. Under a Lorentz transformation Λ, c is unchanged and any expression where the space-time indices of 4-vectors (or more generally tensors) are contracted is Lorentz invariant, e.g. $x'^\mu x'_\mu = x^\mu x_\mu$. So $\Box_x = \Box_{x'}$. Thus, the KG equation $(\hbar^2 \Box_x + m^2 c^2)\psi(x) = 0$ is Lorentz invariant as long as $\psi(x)$ transforms as a scalar under Lorentz transformations $\psi'(x') = \psi(x)$. Since we have not considered any internal (spin) degrees of freedom, the KG equation may be of relevance to spin zero scalar particles, such as pions.

• We also have the covariant and contravariant components of the momentum 4-vector

$$p^\mu = (E/c, \vec{p}) \quad \text{and} \quad p_\mu = \eta_{\mu\nu} p^\nu = (E/c, -\vec{p}) = \left(\frac{i\hbar}{c} \frac{\partial}{\partial t}, i\hbar \nabla \right) = i\hbar \partial_\mu. \tag{20}$$

where we used the correspondence rule, to write the 4-momentum operator in quantum mechanics in terms of the 4-gradient. In terms of the 4-momentum p_μ check that we may write the KG equation as

$$\left(p_\mu p^\mu - m^2 c^2 \right) \psi(x) = 0. \tag{21}$$

3.1.3 Non-relativistic limit

It is possible to obtain the Schrödinger equation in a non-relativistic limit of the KG equation. However, one cannot do this by simply putting $c = \infty$ in the KG equation. Classically, a non-relativistic situation is one where the energy is mostly rest energy. For a free particle, this would mean $E = mc^2 + KE \approx mc^2 + \frac{p^2}{2m}$. In this case, the primary time dependence of a plane wave $\psi(x,t) = e^{i(\vec{k}\cdot x - Et/\hbar)}$ is given by putting $E \approx mc^2$. Of course, there would be some residual time dependence due to the remaining energy. So to facilitate taking the non-relativistic limit, let us change variables to a new wave function $\phi(r,t)$

$$\psi(r,t) = e^{-imc^2 t/\hbar} \phi(r,t) \tag{22}$$

We have in mind that the factor $e^{-imc^2 t/\hbar}$ takes care of the fast time dependence (high frequency) and $\phi(r,t)$ only has a residual slow time dependence. Putting this form in KG, one finds that ϕ satisfies

$$i\hbar \dot{\phi} - \frac{\hbar^2}{2mc^2} \ddot{\phi} = -\frac{\hbar^2}{2m} \nabla^2 \phi. \tag{23}$$

So far we have made no approximation. Now we may take a non-relativistic limit by letting $c \to \infty$, the term second order in time derivatives drops out (ϕ has slow time dependence) and we get the usual free particle SE. An energy eigenstate is then of the form $\phi(r,t) = e^{i(\vec{k}\cdot r - E_{nr} t/\hbar)}$ where $E_{nr} = \hbar^2 \vec{k}^2 / 2m$. Thus for an energy eigenstate, the original wave function is $\psi(r,t) \approx e^{i\vec{k}\cdot\vec{r} - iEt/\hbar}$ where $E = mc^2 + E_{nr}$.

3.1.4 Coupling to electromagnetic field

We can study the KG equation in the presence of an electromagnetic field defined by the scalar and vector potential $\phi, \vec{A}$ in the same way as we did for the Schrödinger equation. We apply

the 'minimal coupling' prescription

$$E \to E - e\phi \quad \text{and} \quad \vec{p} \to \vec{p} - e\vec{A}. \tag{24}$$

to the relativistic energy momentum dispersion relation $E^2 = \vec{p}^2c^2 + m^2c^4$. This is a consistent thing to do since it is shown in electrodynamics that $A_\mu = (\phi/c, -\vec{A})$ transform under Lorentz transformations in the same manner as $p_\mu = (E/c, -\vec{p}) = i\hbar\partial_\mu$, i.e. as the covariant components of a 4-vector. Thus $\pi_\mu = p_\mu - eA_\mu$ is a covariant 4-vector under Lorentz transformations.

• To get a wave equation we then use the correspondence rule $E \to i\hbar\frac{\partial}{\partial t}, \vec{p} \to -i\hbar\nabla$ and treat $\vec{A}$ and ϕ as multiplication operators on the wave function $\psi(x,t)$. The resulting wave equation is

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi\right)^2 \psi = c^2\left(-i\hbar\nabla - e\vec{A}\right)^2 \psi + m^2c^4\psi. \tag{25}$$

This equation can be written in manifestly Lorentz invariant form. Recall that KG could be written as $(p^\mu p_\mu - m^2c^2)\psi = 0$. Coupling to an electromagnetic field simply means we replace $p_\mu \to \pi_\mu = p_\mu - eA_\mu$. The resulting equation is $(\pi^\mu\pi_\mu - m^2c^2)\psi = 0$. Check that this is the same as the above equation. We will say more about the electromagnetic interaction of a relativistic particle when we discuss the Dirac equation.

3.1.5 Local conservation law and physical interpretation

• Recall that a key feature of the SE that made it acceptable as a quantum mechanical wave equation is its physical probability interpretation: the presence of a positive probability density and a current which together satisfy a local conservation law (continuity equation) $\frac{\partial P}{\partial t} + \nabla\cdot j = 0$. We seek a probability density $P(x,t)$ and current $j(x,t)$ for the KG equation that satisfy a continuity equation and reduce to the known non-relativistic quantities in the appropriate limit. Since the non-relativistic probability density and current

$$P_{nr}(x,t) = |\psi(x,t)|^2 \quad \text{and} \quad j_{nr}(x,t) = \frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right) \tag{26}$$

are bilinear, it is simplest to look for a local conservation law bilinear in ψ. We multiply the KG equation $-\hbar^2\ddot{\psi} = -\hbar^2c^2\nabla^2\psi + m^2c^4\psi$ by ψ^* and the complex conjugate equation by ψ and subtract the two to get

$$\frac{\partial}{\partial t}\left(\psi^*\dot{\psi} - \psi\dot{\psi}^*\right) = c^2\nabla\cdot\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right). \tag{27}$$

To match the Schrodinger probability current, if we define

$$P(x,t) = \frac{i\hbar}{2mc^2}\left(\psi^*\dot{\psi} - \psi\dot{\psi}^*\right) \quad \text{and} \quad j(x,t) = \frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right) \tag{28}$$

then $\partial_t P(x,t) + \nabla\cdot j(x,t) = 0$ is a local conservation law for the KG equation.

• What is more, this continuity equation is Lorentz invariant. Let us define the current density

$$j^\mu = \left(cP, \vec{j}\right) = -\frac{\hbar}{m}\Im\left(\psi^*\frac{1}{c}\frac{\partial\psi}{\partial t}, -\psi^*\nabla\psi\right) \tag{29}$$

Since ψ is a scalar under Lorentz transformations, j^μ transforms in the same manner as the 4-vector $\partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right)$. So j^μ are the contravariant components of a 4-vector. Contracting with the covariant 4-divergence, $\partial_\mu j^\mu$ is a Lorentz invariant quantity, which by the continuity equation must vanish $\partial_\mu j^\mu = 0$.

• We check via the substitution $\psi = e^{-imc^2 t/\hbar}\phi$ that

$$P = \frac{i\hbar}{2mc^2}\left(-\frac{2imc^2}{\hbar}|\phi|^2 + \phi^*\dot{\phi} - \phi\dot{\phi}^*\right) \to |\phi|^2 = P_{nr} \quad \text{and} \quad \vec{j} = \vec{j}_{nr} \tag{30}$$

in the non-relativistic limit $c \to \infty$ where ϕ solves the non-relativistic SE.

• So we have a Lorentz-invariant local conservation law for the KG equation, with the correct non-relativistic limit! But can $P(x,t)$ be interpreted as a probability density? No, since it can be negative. Notice that $P = -\frac{\hbar}{mc^2}\Im\psi^*\dot{\psi}$. Since KG is second order in time, both $\psi(x,0)$ and $\dot{\psi}(x,0)$ may be freely specified as initial conditions. E.g., we could take $\psi(x,0) \in \mathbb{R}$ and $\Im\dot{\psi}(x,0) > 0$. Then $P(x,0)$ would be negative. As another example, let us calculate P for a plane wave solution of KG

$$\psi(x,t) = e^{i(\vec{k}\cdot r - \omega t)} \quad \text{where} \quad \hbar\omega = \pm\sqrt{c^2\hbar^2\vec{k}^2 + m^2c^4}. \tag{31}$$

We find $P(x,t)$ is positive for positive energy plane waves and negative for negative energy plane waves:

$$P(x,t) = \frac{\hbar\omega}{mc^2} = \pm\sqrt{1 + \frac{p^2}{m^2c^2}}. \tag{32}$$

Thus we may *not* interpret $P(x,t)$ as a probability density.

• What is more, $P(x,t)$ is identically zero at all times if the initial conditions $\phi(x,0)$ and $\dot{\phi}(x,0)$ are chosen to be real. A real initial condition is a perfectly legitimate initial condition, and one checks that a real initial wave function remains real at all times under Klein-Gordon evolution. It is hard to understand why the KG equation should assign zero 'probability' to such a wave function.

• The lack of a non-negative P and thus the absence of a probability interpretation arises from the fact that the KG equation is second order in time, unlike the SE which is first order in time. Thus, the KG equation cannot be interpreted as a quantum mechanical wave equation for the probability amplitude of one relativistic particle in the same way as the SE is a qm wave equation for one non-relativistic particle. For this reason, as well as due to the spectrum being unbounded below, the KG equation was discarded as a consistent description of a single relativistic spin zero particle.

• In retrospect, it was unreasonable to expect to find a consistent quantum mechanical relativistic wave equation for one particle (or even any fixed number of such particles). This is because it was found experimentally that if a particle is accelerated to an energy significantly in excess of its rest energy, then by the process of particle production and destruction, an indefinite number of additional particles are produced. The number of particles is not fixed in time. However, even in such a situation, one finds that total electric charge is conserved.

• Later, the Klein-Gordon equation was resurrected by Pauli and Weisskopf (1934), who interpreted P as proportional to electric charge density and j as proportional to charge current and the negative energy solutions could be interpreted in terms of particles of opposite charge. In the

context of scalar π-mesons, $P, \vec{j}$ could be interpreted as isospin charge and current. However, in its new incarnation, the KG equation for a complex-valued wave function $\psi(x,t)$ was not a quantum mechanical wave equation at all, but rather an equation for some sort of charged relativistic 'fluid' (field). A situation in which $\phi(x,t)$ is real is then interpreted as an uncharged fluid! Looked at this way, the KG equation could be incorporated as an ingredient in a larger framework of quantum fields, applicable to the relativistic quantum mechanics of an indefinite number of spin-0 particles, rather than to a single spin-0 particle.

• Even aside from its use in the quantum theory of fields, the KG equation has been successfully applied to calculate relativistic corrections to the spectrum of pionic atoms. These are atoms consisting of π^--mesons bound to a nucleus. X-rays are emitted when the π^- makes a transition between levels of such atoms and the spectra are in good agreement with those computed using the KG equation in a Coulomb potential. Rather than describe pionic atoms, we will describe such relativistic corrections for ordinary atoms when we discuss the Dirac equation for electrons.

3.2 Dirac equation

• Some of the difficulties with the KG equation as a relativistic wave equation for the probability amplitude of a single scalar particle (especially the lack of a positive locally conserved probability density) stem from the fact that it is second order in time. Dirac (1928) looked for a relativistic wave equation that is first order in time and admits a non-negative conserved density which could be interpreted as a probability. The simplest first order equation, which follows by applying the correspondence rule to $E = \sqrt{p^2c^2 + m^2c^4}$, is however not manifestly Lorentz invariant. Indeed it is not of finite order in space-derivatives though it is first order in time. Dirac looked for some other way of 'taking this square-root' so that the equation is first order in space derivatives. This would make it easier to ensure Lorentz invariance. He found a remarkable solution to this problem. However, there is a necessary condition for the consistency of any relativistic wave equation: the wave function must satisfy the KG equation. This would ensure that wave packet solutions in the classical limit obey the energy momentum dispersion relation $E^2 = p^2c^2 + m^2c^4$ which must hold for any relativistic particle in classical mechanics. We will use this consistency condition to obtain Dirac's relativistic wave equation for the spin-half electron. Recall that the Pauli equation for a non-relativistic electron was for a two component wave function. So we expect to need at least a two component wave function to account for two linearly independent spin projections. We seek a wave equation of the form

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \quad \text{for} \quad H \quad \text{hermitian} \tag{33}$$

and linear in momenta. For a free particle, any such linear hamiltonian can be written

$$H = c\vec{\alpha}\cdot\vec{p} + \beta mc^2 = -i\hbar c\vec{\alpha}\cdot\nabla + \beta mc^2 \tag{34}$$

where $\vec{\alpha}$ and β are dimensionless and independent of x, t, p, E. Dimensional analysis implies the constant term must be linear in mass and also fixes the factors of $\hbar$ and c. H is called the Dirac hamiltonian or the Dirac operator. We expect ψ to have at least $N = 2$ components, so $\vec{\alpha}$ and β must be *constant* $N \times N$ hermitian matrices[2]. Since they are constant matrices,

[2]If $N = 1$, the hamiltonian wouldn't be rotation-invariant as there would be a preferred vector $\vec{\alpha}$ involved in its specification.

they will commute with the momentum operator p. Other such matrices already appeared in the hamiltonian of the Pauli equation

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{1}{2m}(\vec{\sigma}\cdot\vec{p})^2\psi. \tag{35}$$

But the Pauli equation is not Lorentz invariant, it is second order in space derivatives and first order in time derivatives. It is a non-relativistic equation but can serve as inspiration for the Dirac equation.

• A solution of the Dirac equation automatically satisfies the second order equation

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = H^2\psi = \left[-\hbar^2c^2\alpha_i\alpha_j\partial_i\partial_j + \beta^2m^2c^4 + mc^3(\alpha_i\beta+\beta\alpha_i)p_i\right]\psi. \tag{36}$$

Comparing with the KG equation

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = \left[-\hbar^2c^2\partial_i\partial_i + m^2c^4\right]\psi \tag{37}$$

we find the consistency conditions for Dirac's matrices

$$\alpha_i^2 = \beta^2 = I, \quad [\alpha_i,\beta]_+ = 0 \quad \text{and} \quad [\alpha_i,\alpha_j]_+ = 0 \quad \text{for} \quad i\neq j. \tag{38}$$

In other words, the hermitian α and β matrices must square to the identity and anti-commute in pairs. So their eigenvalues must be ±1. What is more, we can show that they must be traceless

$$\alpha_i\beta+\beta\alpha_i = 0 \;\Rightarrow\; \alpha_i = -\beta\alpha_i\beta \;\Rightarrow\; \text{tr}\,\alpha_i = -\,\text{tr}\,\beta^2\alpha_i = -\,\text{tr}\,\alpha_i \;\Rightarrow\; \text{tr}\,\alpha_i = 0. \tag{39}$$

Similarly, we show that $\text{tr}\,\beta = 0$. Since they are traceless, they must each have an equal number of $+1$ and -1 eigenvalues, so they must be of even dimension N. $N=2$ is however disallowed since we cannot find four such 2×2 matrices (show this!). The next simplest possibility is $N=4$, and it turns out to be possible to find four 4×4 matrices satisfying the above conditions. This also means that the wave function $\psi(x,t)$ must be a four-component column vector (called a 'Dirac' spinor) rather than a two-component vector (sometimes called a 'Pauli' spinor). The adjoint $\psi^\dagger$ is a four component row vector. While we expected to need two components to describe the two spin projections of a spin half particle, the additional two components are unexpected, but forced upon us by internal consistency.

3.2.1 Dirac's 4×4 representation of $\vec{\alpha},\beta$ matrices

• Just as it is convenient to express the spin operators of a non-relativistic particle in a particular basis (say one where S_z is diagonal), it is convenient to pick a basis to represent $\vec{\alpha},\beta$ as specific numerical matrices to facilitate working out explicit solutions of the Dirac equation etc. The representation chosen by Dirac is one where β is diagonal. It facilitates passage to the non-relativistic limit, discussion of spin etc. It is a basis where β is diagonal and its $+1$ eigenvalues precede its -1 eigenvalues:

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{40}$$

is clearly hermitian, traceless and squares to the identity. The condition that α_i anti-commute with β may be used to show that α_i must be block off diagonal. To do so, use hermiticity to write

$$\alpha_i = \begin{pmatrix} a_i & b_i \\ b_i^\dagger & d_i \end{pmatrix} \tag{41}$$

where a_i, d_i, b_i are 2×2 matrices. Then

$$\alpha_i\beta + \beta\alpha_i = \begin{pmatrix} 2a_i & -b_i + b_i \\ b_i^\dagger - b_i^\dagger & -2d_i \end{pmatrix} = 0 \quad \Rightarrow \quad a_i = d_i = 0 \quad \Rightarrow \quad \alpha_i = \begin{pmatrix} 0 & b_i \\ b_i^\dagger & 0 \end{pmatrix}. \tag{42}$$

The conditions $[\alpha_i, \alpha_j]_+ = 2\delta_{ij}$ imply that

$$b_i b_j^\dagger + b_j b_i^\dagger = 0 \quad \text{and} \quad b_i^\dagger b_j + b_j^\dagger b_i = 0 \quad \text{for} \quad i \neq j \quad \text{while} \quad b_i b_i^\dagger = b_i^\dagger b_i = I \quad \text{for} \quad i = 1, 2, 3. \tag{43}$$

The latter conditions say that b_i are unitary matrices. Let us now see if we can identify three matrices b_i with these properties. It is simplest to try to pick b_i to be hermitian. Then the conditions become

$$[b_i, b_j]_+ = 0 \quad \text{for } i \neq j \text{ and} \quad b_i^2 = 1 \quad \text{for} \quad i = 1, 2, 3. \tag{44}$$

So we seek three hermitian and unitary 2×2 matrices that square to the identity and anti-commute in pairs. The Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{45}$$

satisfy these conditions and so we may take $b_i = \sigma_i$. Thus we have the four Dirac matrices

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}. \tag{46}$$

There are other ways to represent the 4×4 Dirac matrices, though they lead to the same physics.

3.2.2 Local conservation law and probability interpretation

• We seek a locally conserved density (hopefully positive) and current for the Dirac equation. By analogy with what worked for the Schrodinger and KG equations, we look for a bilinear conservation law by taking the difference of the Dirac equation and its adjoint after multiplying by $\psi^\dagger$ from the left and ψ from the right respectively. The Dirac equation and its adjoint are

$$i\hbar\frac{\partial\psi}{\partial t} = -i\hbar c\alpha \cdot \nabla\psi + mc^2\beta\psi, \quad \text{and} \quad -i\hbar\frac{\partial\psi^\dagger}{\partial t} = i\hbar c\nabla\psi^\dagger \cdot \alpha + mc^2\psi^\dagger\beta \tag{47}$$

So we get

$$i\hbar\psi^\dagger\dot{\psi} = -i\hbar c\psi^\dagger\alpha \cdot \nabla\psi + mc^2\psi^\dagger\beta\psi \quad \text{and} \quad -i\hbar\dot{\psi}^\dagger\psi = i\hbar c\nabla\psi^\dagger \cdot \alpha\psi + mc^2\psi^\dagger\beta\psi. \tag{48}$$

Subtracting, we get a local conservation law

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) + \nabla \cdot \left(\psi^\dagger c\vec{\alpha}\,\psi\right) = 0 \quad \text{or} \quad \frac{\partial P}{\partial t} + \nabla \cdot \vec{j} = 0 \tag{49}$$

where $P(x,t) = \psi^\dagger\psi$ and $\vec{j}(x,t) = \psi^\dagger c\alpha\,\psi$. Dirac interpreted $P(x,t)$ as a probability density as it is non-negative and j as a probability current density by analogy with Born's interpretation of the Schrödinger wave function. Thus, the problem of negative probabilities in the KG equation could be avoided in the case of the Dirac equation.

3.2.3 Plane waves: Energy and momentum eigenstates of free Dirac Hamiltonian

• Each component of the Dirac wave function $\Psi(\vec{r},t)$ satisfies the KG eqn, which admits plane wave solutions. So it is reasonable to expect the Dirac equation to admit plane waves. We are also interested in energy eigenstates of the Dirac hamiltonian. These are plane waves, as expected of a free particle.

• To find solutions of the Dirac equation, let us proceed by separation of variables. Ψ depends on $\vec{x},t$ as well as spinor degrees of freedom. We make the SOV ansatz

$$\Psi(r,t) = u\psi(r)T(t) \tag{50}$$

where u is a constant (independent of r,t), dimensionless 4-component Dirac spinor. Insertion in the Dirac equation $i\hbar\dot{\Psi} = (c\alpha\cdot p + \beta mc^2)\Psi$ and division by $\psi(r)T(t)$ gives

$$i\hbar u\frac{\dot{T}}{T} = -i\hbar c\frac{\nabla\psi}{\psi}\cdot\vec{\alpha}u + mc^2\beta u = Eu. \tag{51}$$

Lhs is a function of t while rhs a function of r, so both must equal a constant (spinor), which must be proportional to u from the lhs. We denote the proportionality constant E. Thus $T(t)\propto\exp(-iEt/\hbar)$ and

$$-i\hbar c\frac{\nabla\psi}{\psi}\cdot\vec{\alpha}u = (E - mc^2\beta)u \tag{52}$$

Now the rhs is independent of r while the lhs depends on r. So the lhs must be a constant spinor. The dependence of the lhs on r is entirely via the vector $-i\hbar\nabla\psi/\psi$. So this vector must be a constant vector, which we denote $\vec{p}$. Thus $\psi(r)$ and u must satisfy

$$-i\hbar\nabla\psi = \vec{p}\psi \;\text{ or }\; \psi\propto e^{i\vec{p}\cdot r/\hbar} \quad\text{and}\quad \left(c\alpha\cdot p + \beta mc^2\right)u = Eu \tag{53}$$

Thus, separable solutions of the Dirac equation are plane waves

$$\psi(\vec{r},t)\propto u\, e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}. \tag{54}$$

where u is a constant spinor satisfying $\left(c\alpha\cdot p + \beta mc^2\right)u = Eu$. This is a system of four homogeneous linear equations in four unknowns, the components of $u = u(E,\vec{p})$, which are constant in space and time, but could depend on the separation constants E and $\vec{p}$ as well as m and c. For non-trivial solutions to exist, the determinant of the 4×4 matrix of coefficients must vanish.

$$\det\begin{pmatrix}(mc^2-E)I & c\sigma\cdot p\\ c\sigma\cdot p & -(mc^2+E)I\end{pmatrix} = \det\begin{pmatrix} mc^2-E & 0 & cp_3 & c(p_1-ip_2)\\ 0 & mc^2-E & c(p_1+ip_2) & -cp_3\\ cp_3 & c(p_1-ip_2) & -(mc^2+E) & 0\\ c(p_1+ip_2) & -cp_3 & 0 & -(mc^2+E)\end{pmatrix} = 0.$$

This determinant is a quartic polynomial in E namely $(E^2-\vec{p}^2c^2-m^2c^4)^2$ (show this!). Thus, for non-trivial energy and momentum eigenstates to exist, E and $\vec{p}$ must satisfy $(E^2-\vec{p}^2c^2-m^2c^4)^2=0$, so the eigenvalues are $E = E_+, E_+, E_-, E_-$ where $E_\pm = \pm\sqrt{\vec{p}^2c^2+m^2c^4}$.

• Alternatively, we could obtain this relation by recalling that every component of a Dirac wave function must solve the KG equation, and this will be the case for the plane wave only if E and $\vec{p}$ satisfy the above mass-shell condition. We will obtain this condition in yet another way below.

• E is named energy as plane waves are eigenfunctions of the Dirac hamiltonian with eigenvalue E

$$H\Psi = i\hbar\dot{\Psi} = E\Psi. \tag{55}$$

$\vec{p}$ is called momentum since the above plane waves are eigenfunctions of the momentum operator $-i\hbar\nabla\psi = \vec{p}\psi$ with eigenvalue $\vec{p}$. Thus the plane waves are simultaneous eigenfunctions of the energy $i\hbar\frac{\partial}{\partial t}$, momentum $-i\hbar\nabla$ and hamiltonian H operators. This is to be expected, since the free particle hamiltonian is space and time-translation invariant.

• To find the eigenstates and energy spectrum of the Dirac hamiltonian we must find those values of E for which there are non-trivial eigenspinors $u(\vec{p})$ satisfying

$$\begin{pmatrix} mc^2 I & c\sigma\cdot\vec{p} \\ c\sigma\cdot\vec{p} & -mc^2 I \end{pmatrix} u = Eu. \tag{56}$$

We are interested in a massive $m \neq 0$ particle (the electron). The simplest case is $\vec{p} = 0$, when the plane wave does not travel, i.e., the particle is at rest. In this case, the hamiltonian $H = \beta mc^2$ is diagonal and the eigenspinors can be taken as the standard basis spinors

$$u^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \uparrow \\ 0 \end{pmatrix}, \quad u^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \downarrow \\ 0 \end{pmatrix}, \quad u^{(4)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \uparrow \end{pmatrix}, \quad u^{(4)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \downarrow \end{pmatrix}. \tag{57}$$

with energy eigenvalues $E = mc^2, mc^2, -mc^2, -mc^2$. The first two $u^{(1)}, u^{(2)}$ are positive energy eigenspinors with energy equal to that of a particle at rest. The presence of two linearly independent positive energy solutions is to be welcomed, since we wished to model the electron, which is a spin half particle. So without a priori assuming anything about spin, Dirac's formalism, which is based on the relativistic energy momentum dispersion relation, automatically produces an equation for a spin half particle. However, it produces some seemingly unwanted things as well, $u^{(3)}, u^{(4)}$ are negative energy eigenspinors. Despite the Dirac equation being first order in time, negative energy solutions remain. Suitably interpreted, they turn out to be necessary, to accommodate anti-electrons, which are inevitably produced even if a single electron is accelerated to energies much more than its rest energy. So we retain the negative energy solutions, in anticipation of their physical utility.

• More generally, when $\vec{p}$ is not necessarily zero, we suspect that the top two components of u might be significant for positive energy plane waves while the lower components may be significant for negative energy solutions. So it makes sense to write $u = \begin{pmatrix} \phi \\ \chi \end{pmatrix}$, they satisfy the coupled equations

$$\phi = \frac{c\sigma\cdot p}{E - mc^2}\chi \quad \text{and} \quad \chi = \frac{c\sigma\cdot p}{E + mc^2}\phi. \tag{58}$$

Eliminating χ we get

$$c^2(\sigma\cdot\vec{p})^2\phi = (E^2 - m^2c^4)\phi \quad \text{or} \quad (E^2 - m^2c^4 - p^2c^2)I\phi = 0 \tag{59}$$

Similarly we get

$$(E^2 - m^2c^4 - p^2c^2)I\chi = 0. \tag{60}$$

Each is (the same) homogeneous linear equation for a two-component spinor. We have non-trivial solutions provided the determinant of the coefficient matrix $(E^2 - p^2c^2 - m^2c^4)^2$ vanishes. Thus there are two distinct energy eigenvalues and each has multiplicity two

$$E_\pm = \pm\sqrt{\vec{p}^2c^2 + m^2c^4} \tag{61}$$

As the momentum $\vec{p}$ is varied, the spectrum of energies of a Dirac particle extends over the range $(-\infty, -mc^2] \cup [mc^2, \infty)$. This is just as for the Klein Gordon equation, except that here each energy level is twice as degenerate, due to the additional spin degrees of freedom. All energy levels (except $E = \pm mc^2$) are of course infinitely degenerate, as the same energy is obtained irrespective of the direction of momentum.

• Let us obtain the corresponding eigenvectors. Clearly, every two component vector φ solves $(E_+^2 - m^2c^4 - p^2c^2)I\varphi = 0$ and we *could* choose the standard basis $\uparrow = (1,0)^t$ and $\downarrow = (0,1)^t$ for the eigenvectors φ, though any other pair of linearly independent ϕ would also do. The corresponding $\chi's$ are fixed as $\chi = \frac{c\sigma\cdot p}{E+mc^2}\varphi$. Thus we have found two (orthogonal - check this!) eigenspinors corresponding to the positive energy eigenvalue $E = E_+$ (check that $(c\alpha \cdot p + \beta mc^2)u^{(1,2)} = E_+u^{(1,2)}$.)

$$u^{(1)} = \begin{pmatrix} \uparrow \\ \frac{c\sigma\cdot p}{E_+ + mc^2} \uparrow \end{pmatrix} \quad \text{and} \quad u^{(2)} = \begin{pmatrix} \downarrow \\ \frac{c\sigma\cdot p}{E_+ + mc^2} \downarrow \end{pmatrix} \tag{62}$$

• Similarly, we find two orthogonal negative energy $E = E_-$ eigenspinors

$$u^{(3)} = \begin{pmatrix} \frac{c(\sigma\cdot p)}{E_- - mc^2} \uparrow \\ \uparrow \end{pmatrix} \quad \text{and} \quad u^{(4)} = \begin{pmatrix} \frac{c(\sigma\cdot p)}{E_- - mc^2} \downarrow \\ \downarrow \end{pmatrix} \tag{63}$$

Notice that these eigenspinors reduce to the previously determined expressions in the limit $\vec{p} = 0$ where the particle is at rest.

• Combining, separable solutions of the Dirac equation are plane waves. They are simultaneous eigenstates of energy and momentum, so they may be labelled by the corresponding eigenvalues $E, \vec{p}$ which must however satisfy the mass shell condition $E^2 = p^2c^2 + m^2c^4$. An orthogonal basis for these plane waves is

$$\Psi^{(j)}(r,t) = u^{(j)}e^{i(\vec{p}\cdot\vec{r} - Et)/\hbar} \quad \text{for } j = 1,2,3,4. \tag{64}$$

So we can regard the plane waves as labelled by $\vec{p}$ and the sign of E. The corresponding energies are $E = E_+ = \sqrt{p^2c^2 + m^2c^4}$ for $j = 1,2$ and $E = E_- = -E_+$ for $j = 3,4$. There however remains a two-fold degeneracy even after E and $\vec{p}$ have been specified, which may be traced to two possible spin projections, to be discussed below. The spectrum of energies is unbounded both above and below and continuous except for a gap $(-mc^2, mc^2)$ separating positive from negative energy eigenstates. For the electron, the size of this gap is $2mc^2 = 2 \times 511 = 1022$ KeV. The presence of negative energy solutions and a spectrum unbounded from below mean that the Dirac equation suffers from some of the same problems as the KG equation. More on this later.

3.2.4 Non-relativistic limit of plane waves

• What happens to the above plane waves in the non-relativistic limit? We consider a situation in which the energy eigenvalue is mostly rest energy and the velocity is small compared to the

speed of light, $v \ll c$ or $p/mc \ll 1$. For positive energy solutions (which are our primary interest) this would mean $E = E_+ \approx mc^2$. For negative energy solutions. $E = E_- \approx -mc^2$.

• Recall that the eigenvalue problem (56) for the Dirac spinor $u = (\varphi \ \ \chi)^t$ could be written as

$$\chi = \frac{c\sigma \cdot p}{E + mc^2}\varphi, \quad \varphi = \frac{c\sigma \cdot p}{E - mc^2}\chi \tag{65}$$

For a positive energy solution, We see that the components of χ are suppressed compared to those of φ by a factor of order $p/mc \ll 1$. So χ is called the small component and φ the large component in the non-relativistic limit. In fact, in this limit, the positive energy solutions obtained above tend to the non-relativistic spin wave functions with the lower component playing no role

$$u^{(1)} \to \begin{pmatrix} \uparrow \\ 0 \end{pmatrix} \quad \text{and} \quad u^{(2)} \to \begin{pmatrix} \downarrow \\ 0 \end{pmatrix}. \tag{66}$$

Thus the two degenerate positive energy solutions incorporate the two linearly independent spin projections in the non-relativistic limit. In effect the Dirac spinors for plane waves reduce to Pauli spinors.

• For negative energy solutions, χ dominates over φ in the non-relativistic limit. The degenerate negative energy eigenstates $u^{(3)} \to (0 \ \ \uparrow)^t$ and $u^{(4)} \to (0 \ \ \downarrow)^t$ are again distinguished by their spin projections.

3.2.5 Spin and helicity

• In the non-relativistic limit, degenerate positive energy plane wave solutions of the Dirac equation with fixed momentum $\vec{p}$ differ by their spin projections; ditto for negative energy solutions. This is a reflection of the fact that momentum $\hat{p}$ and spin $\vec{S}$ commute with the non-relativistic free particle hamiltonian and we may use momentum and spin projection to label the different degenerate energy eigenstates ((H, p, S^2, S_z) are commuting observables). This degeneracy persists in the relativistic case and we have seen that energy and momentum commute. We seek an observable that commutes with both energy and momentum and can be used to label the plane wave eigenstates. What about spin? Based on the non-relativistic analogy, it is reasonable to introduce the Dirac spin (vector) operator, the 4×4 matrix (it acts as the identity on translational degrees of freedom)

$$\vec{S} = \frac{1}{2}\hbar\vec{\Sigma} = \frac{1}{2}\hbar \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \tag{67}$$

The components S_x, S_y and S_z obviously satisfy the angular momentum commutation relations and $S^2 = (3/4)\hbar^2 I$ as for a spin half particle, except that here we have a doubling of the degrees of freedom. The component of spin in any direction, such as S_z, S_x, S_y or $\hat{n} \cdot S$ for any unit vector $\hat{n}$ has the eigenvalues $\pm\hbar/2$ each with two-fold degeneracy.

• The free particle Schrödinger hamiltonian $H = \frac{1}{2m}p^2 \otimes I$ was proportional to the identity in spin space. However, even for a free Dirac particle, the hamiltonian $H = c\alpha \cdot p + \beta mc^2$ is not proportional to the identity in spin space, this situation is forced on us by the relativistic energy momentum dispersion relation which implies α and β cannot be proportional to the identity. A consequence is that the components of spin are in general not conserved in time. Indeed, spin

does not commute with the Dirac hamiltonian in general[3]. We find using $[\vec{\sigma}\cdot\vec{p},\vec{\sigma}] = 2i\vec{\sigma}\times\vec{p}$ that

$$[H,\vec{S}] = \frac{\hbar}{2}[c\alpha\cdot p + \beta mc^2, \vec{\Sigma}] = \frac{\hbar}{2}\begin{pmatrix} 0 & c[\sigma\cdot p,\vec{\sigma}] \\ c[\sigma\cdot p,\vec{\sigma}] & \end{pmatrix} = i\hbar c\vec{\alpha}\times\vec{p}. \tag{68}$$

It is not just the cartesian components of spin that do not in general commute with H, but also the component of spin in any direction $\hat{n}$. We find that

$$[H,\hat{n}\cdot\vec{S}] = -i\hbar c\vec{\alpha}\cdot\hat{n}\times\vec{p}. \tag{69}$$

However, the component of spin in the direction of momentum, which is called helicity $h = \hat{p}\cdot S$ does commute with H and is conserved

$$h = \vec{S}\cdot\hat{p} = \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma}\cdot\hat{p} & 0 \\ 0 & \vec{\sigma}\cdot\hat{p} \end{pmatrix} \;\Rightarrow\; [H,\hat{p}\cdot\vec{S}] = 0. \tag{70}$$

The eigenvalues of helicity are $\pm\hbar/2 = \lambda\hbar$ with $\lambda = \pm\frac{1}{2}$. Thus, we may choose a basis for the two-dimensional subspace of positive energy E_+ plane waves with given momentum $\vec{p}$ in which helicity is also diagonal. However, $u^{(1,2,3,4)}$ are not helicity eigenstates in general, but positive energy helicity eigenstates may be obtained from linear combinations of $u^{(1)}$ and $u^{(2)}$ and negative energy helicity eigenstates from linear combinations of $u^{(3)}$ and $u^{(4)}$. The positive energy helicity eigenstates are obtained by choosing φ to be eigenvectors of $\hat{p}\cdot S$. If we use spherical polar coordinates for $\vec{p} = (p,\theta,\phi)$ we recall (from spin in a $\vec{B}$ field!), that the eigenvectors of $\hat{p}\cdot S$ corresponding to eigenvalues $\pm\hbar/2$ are

$$\varphi_+ = \chi_+ = \begin{pmatrix} \cos\frac{1}{2}\theta \\ e^{i\phi}\sin\frac{1}{2}\theta \end{pmatrix} \quad\text{and}\quad \varphi_- = \chi_- = \begin{pmatrix} e^{-i\phi}\sin\frac{1}{2}\theta \\ -\cos\frac{1}{2}\theta \end{pmatrix} \tag{71}$$

Thus the helicity eigenspinors with positive energy $E = E_+$ are

$$\lambda = +\frac{1}{2}:\;\; u_+ = \begin{pmatrix} \varphi_+ \\ \frac{c\sigma\cdot p}{E_+ + mc^2}\varphi_+ \end{pmatrix} \quad\text{and}\quad \lambda = -\frac{1}{2}:\;\; u_- = \begin{pmatrix} \varphi_- \\ \frac{c\sigma\cdot p}{E_+ + mc^2}\varphi_- \end{pmatrix} \tag{72}$$

Similarly, the negative energy helicity eigenspinors are

$$\lambda = +\frac{1}{2}:\;\; v_+ = \begin{pmatrix} \frac{c\sigma\cdot p}{E_- - mc^2}\chi_+ \\ \chi_+ \end{pmatrix} \quad\text{and}\quad \lambda = -\frac{1}{2}:\;\; v_- = \begin{pmatrix} \frac{c\sigma\cdot p}{E_- - mc^2}\chi_- \\ \chi_- \end{pmatrix} \tag{73}$$

Unlike the component of spin in a general direction, its component in the direction of momentum ($\vec{p}$) is in a sense adapted to its own motion. So helicity is a very natural observable in addition to being a conserved quantity.

3.2.6 Dirac equation coupled to an EM field and non-relativistic limit

• In the presence of an external electromagnetic field arising from the potentials $\phi, \vec{A}$, the Dirac equation is modified by the replacements

$$E \to E - e\phi \quad\text{and}\quad \vec{p} \to \vec{\pi} = \vec{p} - e\vec{A} \tag{74}$$

[3]However, the components of spin are conserved in the rest frame where $\vec{p} = 0$.

Thus, for a spin half particle of charge e in an electromagnetic field, we get the wave equation

$$(\hat{E} - e\phi)\psi = c\alpha \cdot (\vec{p} - e\vec{A})\psi + mc^2\beta\psi \quad \text{or} \quad i\hbar\frac{\partial\psi}{\partial t} = \left[c\alpha \cdot (p - eA) + e\phi + \beta mc^2\right]\psi \tag{75}$$

In order to examine the non-relativistic limit, we write $\psi = \begin{pmatrix}\varphi\\ \chi\end{pmatrix}$ in terms of the so-called large and small components and get the equation

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\varphi\\ \chi\end{pmatrix} = c\vec{\sigma}\cdot\vec{\pi}\begin{pmatrix}\chi\\ \varphi\end{pmatrix} + mc^2\begin{pmatrix}\varphi\\ -\chi\end{pmatrix} + e\phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix}. \tag{76}$$

• To study the non-relativistic limit, we concentrate on a situation where the energy (or its expectation value) is positive and mostly rest energy $E = mc^2 + \Delta E$ where $\Delta E \ll mc^2$. Change variables to

$$\begin{pmatrix}\varphi\\ \chi\end{pmatrix} = e^{-imc^2t/\hbar}\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} \tag{77}$$

where we have in mind that $\tilde{\chi}, \tilde{\varphi}$ are relatively slowly varying in time. Moreover, we anticipate that $\tilde{\chi}$ is relatively small and hope to eliminate it and get a self-contained equation for $\tilde{\varphi}$. But first we get the coupled pair of equations

$$\begin{aligned} i\hbar\partial_t\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} &= -2mc^2\begin{pmatrix}0\\ \tilde{\chi}\end{pmatrix} + c\vec{\sigma}\cdot\vec{\pi}\begin{pmatrix}\tilde{\chi}\\ \tilde{\varphi}\end{pmatrix} + e\phi\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} \quad \text{or} \\ i\hbar\frac{\partial\tilde{\varphi}}{\partial t} = c\sigma\cdot\pi\tilde{\chi} + e\phi\tilde{\varphi} \quad &\text{and} \quad i\hbar\frac{\partial\tilde{\chi}}{\partial t} = -2mc^2\tilde{\chi} + c\sigma\cdot\pi\tilde{\varphi} + e\phi\tilde{\chi} \end{aligned} \tag{78}$$

The equation for χ in the non-relativistic approximation becomes

$$2mc^2\tilde{\chi} \approx c\sigma\cdot\pi\tilde{\varphi}, \tag{79}$$

assuming $\tilde{\chi}$ is slowly varying in time and the rest energy mc^2 is much more than the electric potential energy. To understand the slowly varying assumption, imagine we have an energy eigenstate ψ with $E = mc^2 + \Delta E$, then $i\hbar\,\partial_t\tilde{\chi} = \Delta E\;\tilde{\chi}$ which is small compared to the $2mc^2\tilde{\chi}$ term above. Eliminating

$$\tilde{\chi} = \frac{\vec{\sigma}\cdot\vec{\pi}}{2mc}\tilde{\varphi} \quad \text{we get} \quad i\hbar\frac{\partial\tilde{\varphi}}{\partial t} = \left[\frac{(\vec{\sigma}\cdot(\vec{p} - e\vec{A}))^2}{2m} + e\phi\right]\tilde{\varphi} \tag{80}$$

which we recognize as the Pauli equation for the wave function $\tilde{\varphi}(\vec{r}, t)$ of a non-relativistic spin half charged particle interacting with an electromagnetic field. The latter equation predicts the correct spin magnetic moment of the electron. This gives us some confidence in the correctness of the Dirac equation, at least in the non-relativistic limit and in the manner it incorporates spin.

3.2.7 Negative energy states, holes and anti-particles

• The negative energy solutions of the Dirac equation have not admitted any physical interpretation. They are problematic since the energy spectrum is not bounded below. If an electron is really described by the Dirac equation, then it is unstable to radiative decay to indefinitely lower energies, in the process radiating an infinite amount energy.

• To avoid this unobserved instability, Dirac (1929) proposed that the zero energy state (the vacuum) is not the one where all states are empty, but one where the negative energy states are all filled with electrons, one per available state (Pauli exclusion), and the positive energy states are empty. This is as in a multi-electron atom, where the inner shells are all filled with electrons. This vacuum state is called the Dirac vacuum, the filled negative energy states are called the filled Dirac sea. A single electron at rest (say from ionizing a hydrogen atom) would then occupy the positive energy state $E = mc^2, \vec{p} = 0$, and would be stable against radiative decay, since the negative energy states are all filled. Thus the instability problem is addressed. Physically realizable situations are regarded as finite departures from the Dirac vacuum. We specify energy and charge by mentioning their values relative to the Dirac vacuum, which by definition is a zero energy, zero charge state.

• For instance, a radiative excitation (induced by a photon of energy $> 2mc^2$) could promote an electron from the Dirac sea to a positive energy state. In this process, we would have both the excited electron and a hole in the sea, this process is called pair creation. The hole could move around the Dirac sea by exchanging places with one of the electrons there. Since the motion of the hole corresponds to the oppositely directed motion of the electron it displaces, holes behave like particles of positive charge. Thus, holes have energy and momentum and behave like particles of positive charge and positive energy and the same mass as electrons. A hole is called an anti-electron or positron. On the other hand, if we had a hole in the Dirac sea, then an electron in a positive energy state could suffer radiative decay, and fall into the hole. As a result, the hole and electron both vanish leaving the Dirac vacuum along with 2 or more photons that are emitted. This process is called electron positron annihilation. Positrons were experimentally observed in cloud chamber experiments by Anderson (1932) and Blackett (1932). Since positrons have opposite electric charge, they bend in the opposite direction to electrons in a constant magnetic field that is applied across the cloud chamber. An annihilation event is identified by two such tracks in a cloud chamber, which meet at a point and abruptly end. At least two photons are produced, which is understandable as a single photon would not conserve momentum (most easily seen in the c.m. frame of the colliding electron and positron). In pair creation, two oppositely bending tracks start all of a sudden from a point. The photons do not leave a track in the cloud chamber since they are uncharged. e^+ e^- pair production from cosmic ray photons was observed when the gamma ray interacted with a nucleus. The nucleus is necessary, again for conservation of momentum.

• Note that we may not freely extract an infinite amount of energy from this system. Suppose we are in a state with one positron at rest, corresponding to a vacancy at the top of the Dirac sea. This state has energy mc^2 more than the Dirac vacuum and charge $+e$. Now an electron deeper in the sea with momentum $|\vec{p}|$ could move up and fill this vacancy, thereby shifting the positron to a state with higher momentum $|\vec{p}|$. This process costs energy, which has to be supplied from outside. The positron has in effect been accelerated by an external agency. So we see that if the hole moves lower down in the Dirac sea, then this costs energy. Indeed, all states that we can produce from the Dirac vacuum have energy more than that of the Dirac vacuum. The Dirac vacuum is the state of lowest energy. There is no free lunch!

• The process of e^+e^- pair production is reversible, the pair could annihilate giving leaving behind two photos. So there is no entropy production in this process. Alternatively, the process of pair annihilation $e^+e^- \to 2\gamma$ is a transformation from one microscopic state to another microscopic state. The entropy in the initial and final states are equal, there is no violation of

the second law of thermodynamics. Entropy increases when there are more ways of constructing the final macroscopic state in terms of microstates than there are ways of constructing the initial state.

• In this new interpretation, the number of electrons is not conserved, due to the processes of pair creation and annihilation, though electric charge is conserved. The Dirac equation by itself is not adequate to describe a system with an indefinite number of particles. For one thing, our interpretation of $\int \psi^\dagger \psi \, d^3x$ as the conserved total probability of finding an electron in the system, would not be consistent, since the number of electrons is not conserved. A new formalism, allowing for creation and annihilation of particles is needed. This is the framework of quantum field theory, which was developed beginning in the 1930s. In the new formalism, there is no need for a filled Dirac sea or for holes in the sea, one directly deals with creation operators for electrons and positrons. However, the physical picture of a filled Dirac sea is still valuable as an aid to thought.

• There are some situations where the contributions of the negative energy solutions of Dirac's equation can be ignored. For instance, we have seen that in the nearly non-relativistic limit, (where energies involved are small compared to electron rest energy 511 KeV) $pc \ll mc^2$, the Dirac equation reduces to the Pauli equation for the upper two components of the Dirac spinor, plus small relativistic corrections.

• Alternatively, suppose we want to build an electron wave packet localized in space with a width of order $|\vec{r}| \lesssim d$, by superposing plane wave solutions of the Dirac equation. The localization in position implies a certain spread of momenta of plane waves that enter the superposition, of order $|p| \lesssim h/d$. It turns out that for localization in position, one necessarily has to include some negative energy plane waves in the superposition, this is seen by Fourier decomposing a wave packet. However, it can be shown that the amplitudes of the negative energy plane waves become appreciable only when their momenta are of order $p \sim mc$. So, as long as $mc \gg \frac{h}{d}$, the negative frequency components contribute negligibly. This is the condition $d \gg \lambda_{Compton} = h/mc$. Thus, as long as we are studying a system where an electron is localized over a region whose linear dimension is large compared to the electron Compton wave length, we may ignore the effects of the negative energy solutions. This condition is satisfied in most of atomic physics, where the electron is localized roughly within a Bohr radius $.5 \times 10^{-10}$m of the nucleus, which is about 20 times its Compton wavelength $2.4 \times 10^{-12} m$. Note that $h = 4.135 \times 10^{-15}$ eV-s and $m_e = .511$ KeV/c^2. Thus, the predictions of the Dirac equation are expected to be accurate in most of atomic physics.

3.2.8 Lorentz covariance of the Dirac equation

• Consider two frames (observers) related by a Lorentz transformation $x' = \Lambda x$. To implement the principles of relativity, we wish to specify how the wave function in the transformed frame $\psi'(x')$ may be constructed from the wave function used by the original observer $\psi(x)$, so that both observers can describe the same physical state. For the laws of physics to be the same for both observers, we want to know whether the Dirac equation and the law of local conservation of probability can be written in such a way that they take the same form in both frames.

• As a first step towards examining the transformation of Dirac's equation under Lorentz trans-

formations, we introduce new notation. Recall the local conservation law

$$\frac{1}{c}\partial_t\psi^\dagger\psi + \nabla\cdot\psi^\dagger\vec{\alpha}\psi = 0 \quad \text{or} \quad \partial_\mu j^\mu = 0 \quad \text{where} \quad j^\mu = (\psi^\dagger\psi, \psi^\dagger\vec{\alpha}\psi) = \psi^\dagger\beta(\beta, \beta\vec{\alpha})\psi \tag{81}$$

For this equation to be Lorentz invariant, we want to choose the law of transformation $\psi(x) \mapsto \psi'(x')$ in such a way that j^μ transforms as a contra-variant 4-vector. In particular we want $\psi^\dagger\psi$ to transform as the zeroth component of a 4-vector. For this to be the case, ψ could not possibly transform as a scalar, as then $\psi'^\dagger(x')\psi'(x') = \psi^\dagger(x)\psi(x)$ which is not the way the zeroth component of a four vector transforms.

• To find the appropriate transformation law for ψ, it is convenient to define the 'Pauli adjoint spinor' $\bar{\psi} = \psi^\dagger\gamma^0$ and four new Dirac γ-matrices $\gamma^\mu = (\gamma^0, \gamma^1, \gamma^2, \gamma^3)$

$$\gamma^0 = \beta \quad \text{and} \quad \gamma^i = \beta\alpha_i \tag{82}$$

so that the conserved probability density and current density may be written

$$j^\mu = \bar{\psi}\gamma^\mu\psi. \tag{83}$$

$\bar{\psi} = \psi^\dagger\gamma^0$ is interesting since unlike $\psi^\dagger\psi$, $\bar{\psi}\psi$ turns out to be Lorentz invariant. Despite appearances, γ^μ is not a four-vector. The γ-matrices are constant matrices, just like α_i and β. There is only one 4×4 representation of Dirac matrices up to unitary equivalence (change of basis in Dirac spinor space). For simplicity, we use the same basis in all frames and take the same set of γ-matrices in every frame of reference[4]. One checks that they anti-commute in pairs and $(\gamma^0)^2 = -(\gamma^i)^2 = I$. So their anti-commutation relations may be written succinctly in terms of the (inverse) Minkowski metric

$$\{\gamma^\mu, \gamma^\nu\} \equiv [\gamma^\mu, \gamma^\nu]_+ = 2\eta^{\mu\nu} I. \tag{84}$$

The invariance of γ-matrices under Lorentz transformations is consistent with the fact that the Minkowski metric is unchanged under Lorentz transformations. While γ^0 is hermitian, γ^i are anti-hermitian, they are all traceless.

• In Dirac's basis, the γ-matrices are

$$\gamma^0 = \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad \text{and} \quad \gamma^i = \beta\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}. \tag{85}$$

• The advantage of the γ^μ over α_i, β is that now Dirac's equation can be written in terms of $p_\mu = i\hbar\partial_\mu$, whose behaviour under Lorentz transformations is known. The Dirac equation

$$i\hbar\frac{\partial\psi}{\partial t} + c\, i\, \hbar\, \alpha_i\partial_i\psi - \beta mc^2\psi = 0, \tag{86}$$

upon multiplying from the left by the non-singular matrix β/c becomes,

$$i\hbar\left(\gamma^0\partial_0\psi + \gamma^i\partial_i\psi\right) - mc\psi = 0 \quad \text{or} \quad (i\hbar\gamma^\mu\partial_\mu - mc)\,\psi = 0 \quad \text{or} \quad (\gamma^\mu p_\mu - mc)\,\psi = 0 \tag{87}$$

Sometimes the Feynman slash notation $\not{\partial} = \gamma^\mu\partial_\mu$ is used to write the Dirac equation as $(i\hbar\not{\partial} - mc)\psi = 0$. Since γ^μ does not transform as a 4-vector, $\gamma^\mu p_\mu$ is not Lorentz invariant. So

[4]One could choose *different* γ matrices in each frame $\gamma^\mu(\Lambda) = U(\Lambda)\gamma^\mu U(\Lambda)^{-1}$, but this is an unnecessary complication and still would not imply that the γ^μ transform as the components of a four-vector.

far, we only introduced new notation. Contraction of indices does not *in itself* imply that the Dirac equation is Lorentz invariant. Indeed we have seen that $\psi(x)$ cannot transform as a scalar (otherwise $\psi^\dagger\psi$ would not transform as the zeroth component of the current 4-vector), so $mc\psi(x)$ cannot be Lorentz invariant. So Dirac's equation is not Lorentz invariant. The next best thing is Lorentz covariance, i.e. every term transforms in the same way so that the equation takes the same form in frames related by Lorentz transformations. It turns out that the Dirac equation is Lorentz covariant with each term transforming as a Dirac spinor under Lorentz transformations.

• Note that the free particle Dirac equation is manifestly space- and time-translation invariant, as neither x nor t appears explicitly in the hamiltonian, indeed both $\hat{P}$ and $\hat{E}$ commute with $\hat{H}$.

• Now we want to demand that the Dirac equation take the same form in frames related by Lorentz transformations. The question is whether we can prescribe suitable rules for the transformation of the Dirac spinor $\psi(x)$ so that the Dirac equation is Lorentz covariant. Let us make a Lorentz transformation $x' = \Lambda x$ or $x'^\mu = \Lambda^\mu_\nu x^\nu$ so that $x = \Lambda^{-1}x'$ and

$$\frac{\partial x'^\mu}{\partial x^\nu} = \Lambda^\mu_\nu, \quad \frac{\partial x^\nu}{\partial x'^\mu} = (\Lambda^{-1})^\nu_\mu \quad \text{and} \quad \partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = (\Lambda^{-1}{}^\nu_\mu)\partial_\nu. \tag{88}$$

Under this Lorentz transformation $\psi(x) \mapsto \psi'(x')$. Since the Dirac equation is linear we suppose that the new spinor wave function is related to the old one by some non-singular linear transformation so that its linearity may be preserved. Let $S(\Lambda)$ be a 4×4 matrix acting on the old Dirac spinor, then

$$\psi'(x') = S(\Lambda)\psi(x) \quad \text{and} \quad \psi(x) = S(\Lambda)^{-1}\psi'(x'). \tag{89}$$

• If the Dirac equation takes the same form, then we must have in the old and new frames:

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi(x) = 0 \quad \text{and} \quad (i\hbar\gamma^\mu\partial'_\mu - mc)\psi'(x') = 0. \tag{90}$$

Can we choose $S(\Lambda)$ appropriately so that the former implies the latter?

• Writing $\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu_\mu\partial_\nu$ the Dirac equation in the new frame becomes

$$(i\hbar\gamma^\mu(\Lambda^{-1})^\nu_\mu\partial_\nu - mc)S(\Lambda)\psi(x) = 0. \quad \text{or} \quad (i\hbar S(\Lambda)^{-1}\gamma^\mu S(\Lambda)(\Lambda^{-1})^\nu_\mu\partial_\nu - mc)\psi(x) = 0 \tag{91}$$

on multiplying by $S(\Lambda)^{-1}$. For this to be implied by the Dirac equation in the old variables, we need

$$S(\Lambda)^{-1}\gamma^\mu S(\Lambda)(\Lambda^{-1})^\nu_\mu = \gamma^\nu \qquad \text{or} \qquad S(\Lambda)^{-1}\gamma^\mu S(\Lambda) = \Lambda^\mu_\nu\gamma^\nu. \tag{92}$$

The question is whether we can find a Λ-dependent 4×4 matrix $S(\Lambda)$ with this property.

• Given an infinitesimal Lorentz transformation $\Lambda^\mu_\nu \approx \delta^\mu_\nu + \omega^\mu_\nu$ (ω are infinitesimal real parameters) it is possible to show that the matrix that implements the Lorentz transformation on Dirac spinors is[5]

$$S(\Lambda) \approx I - \frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu} \quad \text{where} \quad \sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu] \quad \text{and} \quad S(\Lambda)^{-1} \approx I + \frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}. \tag{93}$$

[5] And $S(\Lambda) = \exp -\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}$ for the corresponding finite transformation.

Let us sketch why this $S(\Lambda)$ satisfies $S^{-1}\gamma^\mu S = \Lambda^\mu_\nu \gamma^\nu$. First, for an infinitesimal LT, $\omega^{\mu\nu}$ are small and we can certainly write $S(\Lambda) \approx I - \frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}$ for some constant 4×4 matrices $\sigma^{\mu\nu}$ with $\mu,\nu = 0,1,2,3$. Moreover, the condition that Λ preserve the metric $\Lambda^\mu_\rho \Lambda^\nu_\sigma \eta_{\mu\nu} = \eta_{\rho\sigma}$ implies that $\omega_{\mu\nu}$ is anti-symmetric

$$(\delta^\mu_\rho + \omega^\mu_\rho)(\delta^\nu_\sigma + \omega^\nu_\sigma)\eta_{\mu\nu} = \eta_{\rho\sigma} \quad \Rightarrow \quad \omega_{\rho\sigma} + \omega_{\sigma\rho} = 0. \tag{94}$$

Thus, the part of $\sigma^{\mu\nu}$ that is symmetric in μ and ν, does not contribute to $S(\Lambda)$ and we may take $\sigma^{\mu\nu}$ to be anti-symmetric in μ,ν. The commutator $[\gamma^\mu,\gamma^\nu]$ is anti-symmetric in μ,ν and it is checked by a direct calculation that $S^{-1}\gamma^\mu S = \Lambda^\mu_\nu\gamma^\nu$ up to terms quadratic in ω. For this, use is made of the identity

$$[\gamma^\alpha,[\gamma^\mu,\gamma^\nu]] = 4(\eta^{\alpha\mu}\gamma^\nu - \eta^{\alpha\nu}\gamma^\mu) \tag{95}$$

• Thus, it can be shown that the Dirac equation takes the same form in all frames related by Lorentz transformations that can be built from infinitesimal ones, provided ψ *transforms as a spinor* under Lorentz transformations:

$$\psi'(x') = S(\Lambda)\psi(x) = e^{-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}}\psi(x) \approx \left(I - \frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu} + \cdots\right)\psi(x). \tag{96}$$

3.2.9 Lorentz invariance of the continuity equation

• Let us now see whether the local conservation law for probability $\partial_\mu j^\mu = 0$ is Lorentz invariant if ψ transforms as a Lorentz spinor

$$\psi'(x') = S(\Lambda)\psi(x) \quad \text{and} \quad \psi'^\dagger(x') = \psi^\dagger(x)\, S(\Lambda)^\dagger. \tag{97}$$

For this we need to check whether $j^\mu(x) = \psi^\dagger(x)\gamma^0\gamma^\mu\psi(x)$ transforms as a four vector.

• The transformed probability and current density are

$$j'^\mu = \psi'^\dagger(x')\gamma^0\gamma^\mu\psi'(x') = \psi^\dagger(x)S^\dagger\gamma^0\gamma^\mu S\psi(x) = \psi^\dagger S^\dagger\gamma^0 S S^{-1}\gamma^\mu S\psi(x) = \psi^\dagger S^\dagger\gamma^0 S\Lambda^\mu_\nu\gamma^\nu\psi(x). \tag{98}$$

where use has been made of $S^{-1}\gamma^\mu S = \Lambda^\mu_\nu\gamma^\nu$. The question is whether $j'^\mu(x') = \Lambda^\mu_\nu j^\nu(x)$?

• For this to be the case, we need to show that[6]

$$S^\dagger\gamma^0 S = \gamma^0 \quad \text{or} \quad \gamma^0 S^\dagger\gamma^0 = S^{-1}. \tag{99}$$

We will show that $S^\dagger\gamma^0 S = \gamma^0$ for infinitesimal L.T. We first observe that as $\omega_{\mu\nu}$ are real,

$$S = I - \frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu} \quad \Rightarrow \quad S^\dagger \approx I + \frac{i}{4}(\sigma^{\mu\nu})^\dagger\omega_{\mu\nu} \quad \text{where} \quad \sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu,\gamma^\nu] \tag{100}$$

It can be checked that σ^{ij} are hermitian while σ^{0i} are anti-hermitian. Thus we can write

$$S^\dagger \approx I + \frac{i}{4}(\sigma^{ij}\omega_{ij} - \sigma^{0i}\omega_{0i} - \sigma^{i0}\omega_{i0}) \quad \text{and} \quad S^{-1} = I + \frac{i}{4}\left(\sigma^{ij}\omega_{ij}\sigma^{0i}\omega_{0i} + \sigma^{i0}\omega_{i0}\right). \tag{101}$$

More over, γ^0 commutes with σ^{ij} while it anti-commutes with σ^{0i} and so also with σ^{i0} (show this!).

[6]Note that in general $S(\Lambda)$ is not unitary. It is unitary for rotations, but not for boosts.

• With these facts, we may now calculate to leading order in ω

$$\begin{aligned}\gamma^0 S^\dagger \gamma^0 &\approx \gamma^0 I + \frac{i}{4}\gamma^0\left[\sigma^{ij}\omega_{ij} - \sigma^{0i}\omega_{0i} - \sigma^{i0}\omega_{i0}\right]\gamma^0 \\ &= \left[\gamma^0 + \frac{i}{4}\left(\sigma^{ij}\omega_{ij}\gamma^0 + \sigma^{0i}\omega_{0i}\gamma^0 + \sigma^{i0}\omega_{i0}\gamma^0\right)\right]\gamma^0 = S^{-1}(\gamma^0)^2 = S^{-1} \qquad (102)\end{aligned}$$

Thus we have shown $\gamma^0 S^\dagger \gamma^0 \approx S^{-1}$ or $S^\dagger \gamma^0 S \approx \gamma^0$ for infinitesimal LT. A similar proof of $S^\dagger \gamma^0 S = \gamma^0$ also works for finite Lorentz transformations where $S(\Lambda) = \exp -\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}$. It follows that the current transforms as a contravariant four vector. Hence the law of conservation of probability $\partial_\mu j^\mu = 0$ is Lorentz invariant!

3.2.10 Dirac equation in a Coulomb potential: relativistic corrections

• The mean speed v/c of an electron (as a fraction of the speed of light) in the g.s. of the hydrogen atom is of order the fine-structure constant $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}$. To see this quickly, use the Bohr model. Let $k = (4\pi\epsilon_0)^{-1}$. Balance between centripetal acceleration and electrostatic force and quantisation of angular momentum give

$$\frac{mv^2}{r} = \frac{kZe^2}{r^2} \quad \text{and} \quad L = mvr = n\hbar \quad \Rightarrow \quad r_n = \frac{n^2\hbar^2}{mkZe^2} \quad \text{and} \quad v_n = \frac{nkZe^2}{\hbar} = ncZ\alpha. \quad (103)$$

So in the ground state of a hydrogenic atom, $\beta = v/c \sim Z\alpha \approx Z/137$. Thus, the effects of c being finite are small (as long as $Z \ll 137$), so we might treat them in perturbation theory. But they could not be ignored, since discrepancies between the non-relativistic spectrum $E_n = -\mathbb{R}/n^2 = -\frac{mc^2\alpha^2}{2n^2}$ and experimental measurements were found. Recall that the energy of an electron in the Bohr atom is its rest energy minus the binding energy

$$E = mc^2 - \frac{mc^2\alpha^2}{2n^2} \qquad (104)$$

We will find the next order correction to this formula, the fine-structure correction of order α^4.

• To find relativistic corrections to the hydrogen spectrum, we study energy eigenstates of the Dirac hamiltonian in the spherically symmetric potential $V(r)$ in the centre of mass frame of the electron-nucleus system. $m = (m_e^{-1} + m_N^{-1})^{-1}$ is the reduced mass of the electron. For a hydrogen atom $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$, but we will work with a general $V(\vec{r})$ to begin with. It is possible to solve the Dirac equation in a Coulomb potential by separation of variables. However, we will not pursue this approach here. Rather we find the leading corrections to the non-relativistic hamiltonian implied by the Dirac equation. These corrections have interesting physical interpretations and lead to the so-called fine-structure of the hydrogen spectrum. The fine structure effects produce corrections of order $\alpha^4 mc^2$ to the Bohr spectrum which is of order $\alpha^2 mc^2$.

• We look for eigenstates with energies a little more than the electron rest energy $\mathcal{E} = mc^2 + E$ where $E \ll mc^2$. Writing the Dirac equation $i\hbar\partial_t\Psi = (c\alpha\cdot p + \beta mc^2 + V)\Psi$ in terms of two-component spinors,

$$\Psi(\vec{r},t) = e^{-i\mathcal{E}t/\hbar}\begin{pmatrix}\varphi(\vec{r})\\ \chi(\vec{r})\end{pmatrix} \quad \Rightarrow \quad E\begin{pmatrix}\varphi\\ \chi\end{pmatrix} = \begin{pmatrix}VI & c\sigma\cdot p\\ c\sigma\cdot p & (V-2mc^2)I\end{pmatrix}\begin{pmatrix}\varphi\\ \chi\end{pmatrix} \qquad (105)$$

we get the system

$$V\varphi + c\sigma\cdot p\,\chi = E\,\varphi \quad \text{and} \quad c\sigma\cdot p\,\varphi - (2mc^2 - V)\chi = E\,\chi. \tag{106}$$

In the non-relativistic limit, the electric potential energy from $V(r)$ as well as E are small compared to mc^2, so we expect χ to be suppressed relative to φ by a factor of order v/c. But rather than make any approximation, we eliminate χ by expressing it in terms of φ

$$\chi = (2mc^2 + E - V(r))^{-1} c(\sigma\cdot p)\varphi \tag{107}$$

Here $(2mc^2+E-V(r))$ is a diagonal operator in position space, and we have its inverse appearing above.

• Thus the self-consistent equation for φ prior to making any approximation is

$$\begin{aligned} c(\sigma\cdot p)(2mc^2 + E - V)^{-1} c\sigma\cdot p\varphi + V\varphi = E\,\varphi \\ \text{or} \quad \frac{1}{2m}(\sigma\cdot p)\left[1 + \frac{E-V}{2mc^2}\right]^{-1}(\sigma\cdot p)\varphi + V\varphi = E\,\varphi. \end{aligned} \tag{108}$$

We treat the electric potential energy and E as small compared $2mc^2$ and expand the inverse in a series keeping only the first two terms to get (the 'hamiltonian' operator H_E below is hermitian)

$$H_E\,\varphi = \left[\left(\frac{p^2}{2m} + V\right) - \frac{1}{2m}(\sigma\cdot p)\frac{(E-V)}{2mc^2}(\sigma\cdot p)\right]\varphi \approx E\,\varphi. \tag{109}$$

In the n.r. limit $c\to\infty$, the second term in H_E drops out and we recover Pauli's equation for two component spinors. The leading departure from the n.r. limit is obtained by analyzing this new term.

The part of this term involving E is written

$$-\frac{1}{4m^2c^2}E(\sigma\cdot p)^2 = -\frac{1}{8m^2c^2}(p^2E + Ep^2). \tag{110}$$

• The part involving the potential is is rewritten using $\vec{p}V(\vec{r}) = V(\vec{r})\vec{p} - i\hbar\vec{\nabla}V = V\vec{p} + (\vec{p}V)$. We have

$$(\sigma\cdot p)V(\sigma\cdot p) = (\vec{\sigma}\cdot(\vec{p}V))(\sigma\cdot p) + Vp^2 \tag{111}$$

Though the lhs is manifestly hermitian, the two terms on the rhs are *not individually* hermitian (though their sum is!). Since these two terms contribute to different physical processes, we would like to write each in a manifestly hermitian form. We may do so by adding the adjoint and dividing by two:

$$(\sigma\cdot p)V(\sigma\cdot p) = \frac{1}{2}\left(Vp^2 + p^2V\right) + \frac{1}{2}\left[(\sigma\cdot(pV))(\sigma\cdot p) - (\sigma\cdot p)(\sigma\cdot(\vec{p}V))\right] \tag{112}$$

Show that the adjoint of $(\sigma\cdot(pV))(\sigma\cdot p)$ is as indicated. Now use $(\sigma\cdot a)(\sigma\cdot b) = a\cdot b + i\sigma\cdot a\times b$ to write this as

$$(\sigma\cdot p)V(\sigma\cdot p) = \frac{1}{2}\left(Vp^2 + p^2V\right) + \frac{1}{2}\left[(pV)\cdot p - p\cdot(pV) + i\sigma\cdot(pV)\times p - i\sigma\cdot p\times(pV)\right]. \tag{113}$$

Furthermore, we simplify

$$\vec{\sigma}\cdot(\vec{p}V)\times\vec{p} - \vec{\sigma}\cdot\vec{p}\times(\vec{p}V) = 2\vec{\sigma}\cdot(\vec{p}V)\times\vec{p} \quad \text{and} \quad (pV)\cdot p - p\cdot(pV) = -(p^2V) = \hbar^2\nabla^2 V. \tag{114}$$

Thus

$$(\sigma\cdot p)V(\sigma\cdot p)=\frac{1}{2}\left(Vp^2+p^2V\right)+\frac{1}{2}\left[\hbar^2\nabla^2V+2i\vec{\sigma}\cdot(\vec{p}V)\times\vec{p}\right]. \tag{115}$$

So the hamiltonian becomes

$$H_E=\frac{p^2}{2m}+V-\frac{1}{8m^2c^2}\left(p^2E+Ep^2-Vp^2-p^2V\right)+\frac{\hbar^2}{8m^2c^2}\nabla^2V+\frac{\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{\nabla}V)\times\vec{p}. \tag{116}$$

Since the 3rd term is suppressed by $1/c^2$ we may approximate $(E-V)\,\varphi$ by its non-relativistic value $\frac{p^2}{2m}\varphi$. So, at leading order beyond the n.r. limit, we have (writing H for H_E as E no longer appears in it)

$$H\varphi=\left[\frac{p^2}{2m}+V(\vec{r})-\frac{p^4}{8m^3c^2}+\frac{\hbar^2}{8m^2c^2}\nabla^2V+\frac{\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{\nabla}V)\times\vec{p}\right]\varphi=E\varphi. \tag{117}$$

We recognize $\frac{p^4}{8m^3c^2}$ as arising from expanding the square-root $(m^2c^4+p^2c^2)^{\frac{1}{2}}=mc^2+\frac{p^2}{2m}-\frac{p^4}{8m^3c^2}+\cdots$. But interestingly, the Dirac equation implies two additional relativistic corrections (at order $1/c^2$), that are absent in this expansion. In particular, the last term is a spin-dependent energy, which is absent in the n.r. hydrogen hamiltonian. Thus, it is possible to distinguish between the relativistic corrections predicted by the Dirac equation from those predicted by the KG equation or its 'square-root' $i\hbar\partial_t\psi=\sqrt{m^2c^4-\hbar^2c^2\nabla^2}\psi$. Of course, the KG equation and its 'square root' do not incorporate spin half and therefore are somewhat inadequate to describe an electron.

• The relativistic corrections predicted by the Dirac equation simplify for a central potential, where $\nabla V=\hat{r}\partial_rV$ and $\nabla V\cdot\nabla\varphi=\partial_rV\partial_r\varphi$ and $\nabla V\times\vec{p}\varphi=r^{-1}\partial_rV(\vec{r}\times\vec{p})\varphi$. Thus eigenstates must satisfy

$$H\varphi=\left[\frac{p^2}{2m}+V(\vec{r})-\frac{p^4}{8m^3c^2}+\frac{\hbar^2}{8m^2c^2}(\nabla^2V)+\frac{1}{2m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}\vec{L}\cdot\vec{S}\right]\varphi\approx E\,\varphi. \tag{118}$$

The last three terms in H, are all suppressed by c^{-2} compared to the non-relativistic hamiltonian. They lead to the so-called fine-structure of the hydrogen spectrum. The first of these $H_{\rm rel}$ may be attributed to the relativistic energy momentum dispersion relation. The last term $H_{\rm SO}$ represents energy due to so-called spin-orbit coupling and the penultimate term is another relativistic correction called the Darwin term H_D.

3.2.11 Fine structure of hydrogen spectrum

• If we treat these relativistic corrections as perturbations to $H_0=\frac{p^2}{2m}+V(r)$, then in first order perturbation theory, they lead to shifts in the energies of unperturbed eigenstates. The Bohr spectrum is of course highly degenerate, for fixed principal quantum number n, there are $2n^2$ linearly independent states degenerate in energy $-\mathbb{R}/n^2$. The first order correction to energies in degenerate perturbation theory is given by the eigenvalues of the matrix of the perturbing hamiltonian within the degenerate subspace. To find the matrix elements of the perturbing hamiltonians $H_D, H_{\rm SO}, H_{\rm rel}$ within the degenerate subspace of fixed n, we must pick a basis for that subspace. A basis for the degenerate subspace can be chosen as common eigenstates of the pairwise commuting operators H_0, L^2, L_z, S^2, S_z leading to the labels $n, l, m_l, s=\frac{1}{2}, m_s$

corresponding to the 'uncoupled' basis. But there is nothing sacred about this basis. We could also choose within a degenerate energy eigenspace a basis in which H_0, J^2, J_z, L^2, S^2 are diagonal, leading to the labels njm_jls.

• For hydrogen, $V(r) = -ke^2/r$ where $k = 1/4\pi\epsilon_0$. Since $\nabla^2 \frac{1}{r} = -4\pi\delta^3(\vec{r})$ we find

$$H_{\rm Darwin} = \frac{\hbar^2}{8m^2c^2}(\nabla^2 V) = \frac{\pi k e^2 \hbar^2}{2m^2c^2}\delta^3(\vec{r}). \tag{119}$$

Thus the H_D represents a point-like repulsion at the origin. H_D is a spherically symmetric perturbing potential, and it commutes with L^2, L_z, S^2, S_z. So H_D is in fact diagonal in the uncoupled basis for a degenerate subspace of fixed n: $\langle nl'm_l'm_s'|H_D|nlm_lm_s\rangle \propto \delta_{ll'}\delta_{m_lm_l'}\delta_{m_sm_s'}$. Its diagonal matrix elements are its eigenvalues and they are the 1st order shifts to the unperturbed energies. Since $H_D \propto \delta^3(\vec{r})$ the expectation value of H_D vanishes in an unperturbed state whose wave function vanishes at the origin. Since $\psi_{nlm} \propto r^l e^{-r/na_0}$, the Darwin term can only affect the energies of S-wave states. So for $l = 0$,

$$\Delta E_D = \frac{\pi k e^2 \hbar^2}{2m^2c^2}\langle\psi_{n00}|\delta^3(\vec{r})|\psi_{n00}\rangle = \frac{\pi e^2\hbar^2}{2m^2c^2(4\pi\epsilon_0)}|\psi_{n00}(\vec{0})|^2 \tag{120}$$

Moreover, from tabulated hydrogen wave functions, $|\psi_{n00}(\vec{0})|^2 = |R_{n0}(0)|^2/4\pi = 1/(\pi a_0^3 n^3)$. Thus

$$\Delta E_D = \frac{m(ke^2)^4}{2\hbar^4c^2n^3} = \frac{1}{2}mc^2\alpha^4\frac{1}{n^3}\delta_{l0} \qquad \text{as} \qquad a_0 = \frac{\hbar^2}{mke^2}. \tag{121}$$

• The perturbation due to relativistic dispersion relation $H_{\rm rel} = -\frac{p^4}{8m^3c^2}$ also commutes with L^2, L_z, S^2, S_z so it is diagonal within each degenerate subspace of H_0. At first order in P.T. it leads to a correction in energies of $\Delta E_{\rm rel} = \langle\psi|H_{\rm rel}|\psi\rangle$ where ψ_{nlm} are the normalized eigenfunctions of the n.r. hydrogen atom. Using hermiticity of p^2 and the fact that $\left(\frac{p^2}{2m} + V\right)\psi = E_n\psi$ where $E_n = -\mathbb{R}/n^2$ for the unperturbed hydrogen eigenstates, we have the first order correction

$$\begin{aligned}\Delta E_{\rm rel} &= -\frac{1}{8m^3c^2}\langle p^2\psi|p^2\psi\rangle = -\frac{1}{2mc^2}\langle\psi|(E-V)^2|\psi\rangle = -\frac{1}{2mc^2}\left[E_n^2 + \langle V^2\rangle - 2E_n\langle V\rangle\right] \\ &= -\frac{1}{2mc^2}\left[E_n^2 + k^2e^4\langle r^{-2}\rangle + 2E_nke^2\langle r^{-1}\rangle\right] \quad \text{where} \quad k = \frac{1}{4\pi\epsilon_0}.\end{aligned} \tag{122}$$

The expectation value of $1/r$ and $1/r^2$ in hydrogen eigenstates can be calculated (the former can also be obtained in a quick and dirty manner from the Bohr model[7]) and the results are (see Liboff or Griffiths)

$$\langle r^{-1}\rangle = \frac{1}{n^2a_0} \quad \text{and} \quad \langle r^{-2}\rangle = \frac{1}{n^3\left(l+\frac{1}{2}\right)a_0^2} \tag{123}$$

The resulting relativistic-p^4 correction to energies at 1st order in PT is (use $a_0 = \frac{\hbar^2}{mke^2}$, $E_n = -\frac{mc^2\alpha^2}{2n^2}$)

$$\Delta E_{\rm rel} = -\frac{1}{2mc^2}\left[E_n^2 + \frac{2E_nmk^2e^4}{n^2\hbar^2} + \frac{(mk^2e^4)^2}{(l+\frac{1}{2})n^3\hbar^4}\right] = -\frac{E_n^2}{2mc^2}\left[\frac{4n}{l+\frac{1}{2}} - 3\right] = -\frac{mc^2\alpha^4}{8n^4}\left[\frac{4n}{l+\frac{1}{2}} - 3\right] \tag{124}$$

[7] $E = T + V = \frac{1}{2}mv^2 - ke^2/r$, $mv^2/r = ke^2/r^2$ and $mvr = n\hbar$ give $V_n = -2T_n$. So $-ke^2\langle r^{-1}\rangle = -ke^2/n^2a_0$.

Thus, the relativistic correction to energy levels is down by a factor of $\alpha^2 \sim (1/137)^2 \approx 5\times 10^{-5}$ compared to the unperturbed energies. So the relativistic corrections are of order 10^{-4} – 10^{-5} eV.

• For the Coulomb potential, the perturbation due to spin-orbit coupling is[8]

$$H_{\rm SO} = \frac{ke^2}{2m^2c^2}\frac{1}{r^3}\vec{L}\cdot\vec{S}. \tag{125}$$

$H_0 = \frac{p^2}{2m} + V(r)$ and L^2, L_z, S^2, S_z are simultaneously diagonalizable. But $L\cdot S$ does not commute with L_z nor S_z though it does commute with any function of the radial coordinate. So m_l, m_s are no-longer good quantum numbers in the presence of spin orbit coupling. However, $\vec{J} = \vec{L} + \vec{S}$ does commute with L^2, S^2 and $\vec{L}\cdot\vec{S}$. E.g. $[\vec{J}, L^2] = [\vec{L}, L^2] = 0$ and

$$[(L+S)_i, L\cdot S] = [L_i, L_jS_j] + [S_i, L_jS_j] = i\hbar\epsilon_{ijk}L_kS_j + i\hbar\epsilon_{ijk}L_jS_k = -i\hbar\epsilon_{ijk}L_jS_k + i\hbar\epsilon_{ijk}L_jS_k = 0. \tag{126}$$

So $H_{\rm SO}$ is diagonal in the simultaneous eigenbasis of J^2, J_z, L^2 and S^2. So we use the coupled basis $|n, j, l, s, m_j\rangle$ instead of the uncoupled one $|n, l, m_l, s, m_s\rangle$. In the coupled basis, both H_0 and $H_{\rm SO}$ are diagonal. So the shifts in energy due to spin-orbit coupling is given by the expectation value of $H_{\rm SO}$ in state $|njm_jls\rangle$

$$L\cdot S|njlm_j\rangle = \frac{\hbar^2}{2}\left(j(j+1) - l(l+1) - \frac{3}{4}\right)|njlm_j\rangle. \tag{127}$$

$L\cdot S$ has eigenvalue zero for S-wave states so there is no spin-orbit correction to the energy for $l = 0$. Moreover, the expectation value of $1/r^3$ in the same unperturbed eigenstates is

$$\left\langle njlm_j\left|\frac{1}{r^3}\right|njlm_j\right\rangle = \frac{1}{l(l+\frac{1}{2})(l+1)n^3a_0^3}, \quad \text{for} \quad l \neq 0. \tag{128}$$

Thus the spin-orbit correction to energies at first order in perturbation theory is ($a_0 = \frac{\hbar^2}{mke^2}$)

$$\begin{aligned}\Delta E_{\rm SO} &= \frac{ke^2}{2m^2c^2}\frac{\hbar^2[j(j+1) - l(l+1) - 3/4]}{2l(l+\frac{1}{2})(l+1)n^3a_0^3} = \frac{E_n^2}{mc^2}\frac{n[j(j+1) - l(l+1) - 3/4]}{l(l+\frac{1}{2})(l+1)} \\ &= \frac{1}{4}\alpha^4mc^2\frac{[j(j+1) - l(l+1) - 3/4]}{n^3l(l+\frac{1}{2})(l+1)} \quad \text{for} \quad l \neq 0.\end{aligned} \tag{129}$$

The spin-orbit correction vanishes for $l = 0$.

• Let us collect our results so far

$$\Delta E_{\rm rel} = -\frac{mc^2\alpha^4}{8n^4}\left[\frac{4n}{l+\frac{1}{2}} - 3\right],$$

[8]The spin-orbit energy can be motivated by a classical model, by considering the magnetic dipole energy $H = -\mu\cdot B$ of the electron spin in the magnetic field $\vec{B}$ produced by the proton. In the electron rest frame, the proton goes round it uniformly in a horizontal circle, producing in effect a circular current loop of radius r and current $I = e/T$ where e is the proton charge and T is the period. The magnetic field so produced at the electron is $B = \mu_0 I/2r = \mu_0 e/2rT$ pointing vertically upwards. This B is proportional to the angular momentum of the electron, in the rest frame of the proton, which also points vertically, $\vec{L} = mvr = 2\pi mr^2/T$. So $\vec{B} = \frac{\mu_0 e\vec{L}}{4\pi mr^3} = \frac{ke\vec{L}}{mc^2r^3}$. On the other hand, the electron spin magnetic moment is $\vec{\mu} = (-eg/2m)\vec{S}$ where $g \approx 2$. Combining these one gets a spin-orbit energy $H = -\mu\cdot B = \frac{ke^2}{m^2c^2r^3}\vec{L}\cdot\vec{S}$. The result from this simple-minded calculation is twice as big as that obtained from the Dirac equation. The discrepancy was explained by Thomas and is due to fact that the electron's rest frame is not an inertial frame compared to the proton rest frame.

$$\begin{aligned}\Delta E_D &= \frac{1}{2}mc^2\alpha^4\frac{1}{n^3}\delta_{l0},\\ \Delta E_{\rm SO} &= (1-\delta_{l0})\frac{1}{4}\alpha^4 mc^2\frac{[j(j+1)-l(l+1)-3/4]}{n^3 l(l+\frac{1}{2})(l+1)}.\end{aligned} \tag{130}$$

The fine-structure correction is the sum of these three. For $l=0$ only ΔE_D and $\Delta E_{\rm rel}$ contribute, so

$$\Delta E^{l=0} = \Delta E_D + \Delta E_{\rm rel} = -\frac{1}{2}mc^2\frac{\alpha^4}{n^2}\left[\frac{1}{n^2}\left(2n-\frac{3}{4}\right)-\frac{1}{n}\right] = -\frac{1}{2}mc^2\frac{\alpha^2}{n^2}\frac{\alpha^2}{n^2}\left[n-\frac{3}{4}\right]. \tag{131}$$

For $l>0$ only $\Delta E_{\rm SO}$ and $\Delta E_{\rm rel}$ contribute, so

$$\Delta E^{l>0} = \Delta E_{\rm rel}+\Delta E_{\rm SO} = \frac{mc^2\alpha^4}{2n^4}\left[\frac{3}{4}-n\left\{\frac{1}{l+\frac{1}{2}}-\frac{j(j+1)-l(l+1)-3/4}{2l(l+\frac{1}{2})(l+1)}\right\}\right] = \frac{1}{2}\frac{mc^2\alpha^4}{n^4}\left[\frac{3}{4}-\frac{n}{j+\frac{1}{2}}\right].$$

The last equality follows by noting that $\vec{J}=\vec{L}+\vec{S}$ and $s=\frac{1}{2}$. By the rules for addition of angular momentum, $j=l\pm\frac{1}{2}$ for $l>0$. Consequently, it is possible to eliminate l in favor of j and one finds the remarkably simple expression given. What is more, this formula reduces to the previous expression for $\Delta E^{l=0}$ when $l=0$.

• Combining we get a common formula for the hydrogen spectrum including fine structure

$$E_{n,j} = mc^2 - \frac{mc^2\alpha^2}{2n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{j+\frac{1}{2}}-\frac{3}{4}\right)+\cdots\right]. \tag{132}$$

• The energy eigenstates of hydrogen, after including first order effects of spin-orbit coupling, relativistic p^4-correction and Darwin term, may be labelled by the good quantum numbers n,j,m_j,l,s, while m_l,m_s are no-longer good quantum numbers. However, the energies depend only on n and j. So for fixed n and j, states with different values of m_j and l are degenerate in energy. For fixed n, $l=0,1,\cdots,n-1$, $j=l-\frac{1}{2},l+\frac{1}{2}$ (except when $l=0$ when $j=\frac{1}{2}$), $m_j=-j,-j+1,\cdots,j-1,j$. The non-relativistic degeneracy among states with a common value of n is partly lifted by relativistic effects. This is called fine structure splitting and its magnitude is controlled by α, which was called the fine-structure constant by Sommerfeld. For fixed 'n', the n levels $j=\frac{1}{2},3/2,\cdots,n-\frac{1}{2}$ form a so-called fine structure multiplet. Since the fine-structure correction is negative definite (as the smallest possible value of $n/(j+\frac{1}{2})$ is 1), the net effect of relativistic corrections is to increase the binding energy compared to what one expects based on a non-relativistic treatment.

• It is conventional, following the non-relativistic spectroscopic notation, to denote the energy levels by specifying n,l and j in the form nL_j where L is the letter $S,P,D,F,G\ldots$ standing for[9] $l=0,1,2,3,4\ldots$ e.t.c. So the g.s. is $1S_{\frac{1}{2}}$. m_j is not explicitly indicated.

• In general, within a fine structure multiplet, (fixed n), states with higher j have higher energy (less binding energy). The spectroscopic notation for low-lying hydrogen energy levels are given below in increasing order of energy, with degeneracy indicated by equality.

$$1S_{\frac{1}{2}} \;<\; 2S_{\frac{1}{2}}=2P_{\frac{1}{2}} \;<\; 2P_{3/2} \;<\; 3S_{\frac{1}{2}}=3P_{\frac{1}{2}} \;<\; 3P_{3/2}=3D_{3/2} \;<\; 3D_{5/2},\ldots \tag{133}$$

These energy levels are additionally degenerate since for each n,l,j, there are $2j+1$ linearly independent degenerate states corresponding to distinct values of m_j. For example, the g.s.

[9] S= sharp, P = principal, D = diffuse, F = fundamental denote the originating level in a spectral emission.

$1S_{\frac{1}{2}}$ is doubly degenerate corresponding to $m_j = \pm\frac{1}{2}$. This degeneracy can be broken by an external $\vec{B}$ field (Zeeman effect).

• The fine structure splitting within the $n = 2$ multiplet is $\Delta E = E(2P_{3/2}) - E(2S_{1/2}) = \frac{mc^2\alpha^4}{32} = 4.5 \times 10^{-5}$ eV. A transition between $2P_{3/2}$ and $2S_{1/2}$ therefore corresponds to a spectral line of wave length $\lambda = hc/\Delta E = 2.8$ cm or a frequency 10.9 GHz corresponding to radio or radar waves.

• It is possible to get the hydrogen bound state spectrum by solving the Dirac equation without expanding around the non-relativistic limit. The result is the same as that obtained by Sommerfeld (1916) using the Bohr-Sommerfeld quantization conditions in the old quantum theory

$$E = mc^2 \left[1 + \frac{\alpha^2}{\left(n_r + \sqrt{n_\phi^2 - \alpha^2} \right)^2} \right]^{-1/2} \tag{134}$$

where the 'azimuthal' quantum number $n_\phi = (j + \frac{1}{2})$ and the 'radial' quantum number $n_r = n - n_\phi = n - (j + \frac{1}{2})$. When expanded in powers of α for small α, the first three terms reproduce the rest energy, Bohr spectrum and fine structure corrections obtained above. The fine structure corrections are in good agreement with experimental measurements and there was no known discrepancy till experiments by Lamb and Retherford (1948) showed that the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ levels were not degenerate, this is called the Lamb shift, it is not accounted for by the Dirac equation. It required a quantum field theory of the electromagnetic field to explain the Lamb shift.