PHY 510 Relativistic Quantum Mechanics, U. of Rochester

S. G. Rajeev

May 1, 2007

Contents

Chapter 1

Introduction

1.1 **Relativistic Quantum Mechanics is the combination of Special Relativity and Quantum Mechanics**

1.2 **It leads inevitably to the quantum field theory: quantum systems with an infinite number of degrees of freedom.**

1.3 **In another direction, it led to the study of unitary representations of the Lorentz group, and to Harish-Chandra's marvellous theory on the unitary dual of semi-simple Lie groups.**

1.4 **The central object of our study will be the Dirac equation.**

Chapter 2

The Axioms of Quantum Mechanics

2.1 Observables of a physical system are represented by hermitean (more precisely self-adjoint) operators on a complex Hilbert space $\mathcal{H}$.

2.2 The eigenvalues of such an operator (which are necessarily real numbers) are the possible outcomes of measuring it.

2.3 States of a physical system are described vectors (more precisely rays) in the Hilbert space.

2.4 If the system is in state $|\psi>\in\mathcal{H}$, the probability of obtaining the value a during a measurement of the observable A is $\frac{|<a|\psi>|^2}{|<\psi|\psi>|^2}$.

2.4.1 Here, $A|a>=a|a>$ so that $|a>$ is the eigenstate (assumed to be unique) of eigenvalue a .

2.5 There is a self-adjoint operator H , the *hamiltonian*, which describes the time evolution of a state

$$i\hbar\frac{\partial}{\partial t}|\psi>=H|\psi>.$$

2.5.1 This is the *Schrödinger* equation.

Chapter 3

Rotations

3.1 The distance between two points with co-ordinates $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ is given by $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$.

3.2 If we translate both vectors by the same amount $x \to x + a$ and $y \to y + a$, the distance is unchanged.

3.3 Similarly if we rotate both the same way, the distance is unchanged.

3.4 The length of a vector is $\sqrt{x^T x}$ where x is thought of as a 1×3 matrix; i.e., a column vector.

3.4.1 x^T stands for the transpose, which is a 1×3 'matrix'; i.e., a row vector.

3.5 A rotation is described by a 3×3 matrix R :

$$x \to Rx.$$

To preserve the length $(Rx)^T(Rx) = x^T x$; i.e., $x^T(R^T R)x = x^T x$.

3.6 Thus a rotation must satisfy

$$R^T R = 1.$$

A matrix satisfying such a condition is called an *orthogonal matrix.*

3.6.1 Not all orthogonal matrices describe rotations. For example $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is orthogonal; it reflects the first component while leaving the others unchanged. This is not possible by any rotation. On the other hand $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ reflects the first and second components; this just a rotation around the third axis by 180^o .

3.7 A rotation is an orthogonal matrix whose determinant is one.

3.7.1 Since $\det AB = \det A \det B$ and $\det A^T = \det A$, we can deduce that for and orthogonal matrix $(\det R)^2 = 1$. Thus $\det R = \pm 1$. Under small changes of the matrix elements of R , the determinant cannot change: it would have to jump from 1 to -1 if it were to change. Since all rotations can be got from the identity by a continuous change of matrix elements (change the angle of rotation), they have to have determinant one. An orthogonal matrix with determinant -1 is a combination of a rotation and a reflection.

3.8 The set of orthogonal matrices is a *group*

3.8.1 The product of two orthogonal matrices is orthogonal, the identity is an orthogonal matrix and the inverse of an orthogonal matrix is one as well. Moreover the multiplication of matrices is associative.

3.8.2 This group is denoted by $O(3)$

3.9 The set of Special Orthogonal matrices, $SO(3)$, which represent rotations, is a subgroup.

3.10 An infinitesimal transformation $R = 1 + A$ is orthogonal if $A^T + A = 0$; i.e., infinitesimal rotations are described by anti-symmetric matrices.

3.11 An arbitrary anti-symmetric matrix can be written as a linear combination of the basic ones

$$S_{12} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad S_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad S_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

3.12 A *Lie algebra* is a linear vector space along with a bilinear operation satisfying

$$[A, A] = 0, \quad [A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

3.13 The *commutator* of two matrices is defined to be $[A, B] = AB - BA$.

3.14 The set of anti-symmetric matrices form a Lie algebra, using the commutator.

3.14.1 The commutation relations of the basis elements describe this algebra completely:

$$[S_{12}, S_{13}] = -S_{23}, \quad [S_{12}, S_{23}] = S_{13}, \quad [S_{13}, S_{23}] = -S_{12}$$

3.14.2 It is special to the case $n = 3$ that $SO(n)$ has dimension n ; in general the dimension is $\frac{n(n-1)}{2}$. It is convenient to take advantage of this coincidence and use a simplified notation

$$S_3 = S_{12}, S_1 = S_{23}, \quad S_2 = -S_{13}$$

which satisfy

$$[S_3, S_2] = S_1, \quad [S_3, S_1] = -S_2, \quad [S_1, S_2] = S_3$$

3.15 What is the Lie algebra of rotations in n dimensions?

Chapter 4

Spinors

4.1 A linear operator $U : \mathcal{H} \to \mathcal{H}$ in a complex Hilbert space is *Unitary* if it has an inverse and if it preserves the length of all vectors.

4.1.1 In other words, $< U\psi, U\psi > = < \psi, \psi >$ or $U^\dagger U = 1 = UU^\dagger$.

4.2 The set of unitary operators on a complex Hilbert space of dimension n forms a group, the *Unitary group* $U(n)$.

4.2.1 The determinant of a unitary matrix is a complex number of modulus one.

4.2.2 The subgroup of operators which are of determinant one as well is $SU(n)$.

4.2.3 In the simplest case $n = 1$, $U(1)$ the group of complex numbers of modulus one.

4.2.4 When $n = 2$, the general element of $SU(2)$ can be written as $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfying

$$|a|^2 + |c|^2 = 1 = |b|^2 + |d|^2, \quad a^* b + c^* d = 0, \quad ad - bc = 1.$$

In fact we can eliminate c, d from these equations to get a nice description of the elements of $SU(2)$:

$$g = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad |a|^2 + |b|^2 = 1.$$

Thus the points in $SU(2)$ are in one-one correspondence with points on the three dimensional sphere $a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1$

4.3 A unitary matrix infinitesimally close to the identity is of the form $g = 1 + A$ where

$$A + A^\dagger = 0.$$

4.3.1 That is, A is anti-hermitean; this means $A = iH$ where $H = H^\dagger$ is hermitean.

4.4 If $1 + A \in SU(n)$ and A is infinitesimally small, we have in addition $\mathrm{tr}\ A = 0$.

4.4.1 The set of anti-hermitean matrices of zero trace forms a Lie algebra. The commutator is still of zero trace and anti-hermitean:

$$\mathrm{tr}\ [A, B] = 0, \quad [A, B]^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B].$$

Also, we can verify that the conditions for a Lie algebra are satisfied:

$$[A, A] = 0, \quad [A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

4.4.2 Note that in spite of the appearance of complex numbers, the set of anti-hermitean matrices is a vector space over the *real* numbers: multiplying by i will turn a hermitean matrix into a hermitean one. Watch for these little factors of i , they will cause much grief otherwise!

4.4.3 Hermitean matrices do not form a Lie algebra with respect to the commutator: the commutator of two hermitean matrices is *anti*-hermitean.

4.5 The most general traceless hermitean matrix is $\begin{pmatrix} a & b_1 + ib_2 \\ b_1 - ib_2 & -a \end{pmatrix}$ where a, b_1, b_2 are real.

4.6 The Pauli matrices are defined to be

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

4.6.1 These provide a basis for the space of traceless hermitean 2×2 matrices:

$$\begin{pmatrix} a & b_1 - ib_2 \\ b_1 + ib_2 & -a \end{pmatrix} = a\sigma_3 + b_1\sigma_1 + b_2\sigma_2$$

4.6.2 An anti-hermitean matrix is obtained by multiplying a hermitean matrix by i . Thus $i\sigma_1, i\sigma_2, i\sigma_3$ form a basis for the Lie algebra of $SU(2)$.

4.7 The Pauli matrices satisfy the relations

$$\sigma_1\sigma_2 = i\sigma_3 = -\sigma_2\sigma_1, \quad \sigma_2\sigma_3 = i\sigma_1 = -\sigma_3\sigma_2, \quad \sigma_3\sigma_1 = i\sigma_2 = -\sigma_1\sigma_2,$$

4.8 The Lie algebra of traceless anti-hermitean matrices can thus be expressed in terms of the commutation relations of the Pauli matrices:

$$[-\frac{i}{2}\sigma_1, -\frac{i}{2}\sigma_2] = -\frac{i}{2}\sigma_3, \quad [-\frac{i}{2}\sigma_2, -\frac{i}{2}\sigma_3] = -\frac{i}{2}\sigma_1, \quad [-\frac{i}{2}\sigma_3, -\frac{i}{2}\sigma_1] = -\frac{i}{2}\sigma_2,$$

4.8.1 We could also have written this as

$$[\sigma_1, \sigma_2] = 2i\sigma_3$$

etc. But sneaking in the factor of $-\frac{i}{2}$ displays a remarkable relationship of the Lie algebra of 2×2 traceless anti-hermitean matrices to the Lie algebra of 3×3 anti-symmetric matrices. With

$$-\frac{i}{2}\sigma_1 \to S_1, \quad -\frac{i}{2}\sigma_2 \to S_2, \quad -\frac{i}{2}\sigma_3 \to S_3,$$

we get the same commutation relations. They have the same structure in spite of the fact that they are made of different kinds of matrices.

4.9 The Lie algebras of $SU(2)$ and $SO(3)$ are isomorphic.

4.9.1 This rather peculiar mathematical fact is useful in quantum mechanics; we can use it to describe the spin of an electron.

4.10 The operators describing the intrinsic (or *spin*) angular momentum components of an electron at rest are,

$$\frac{\hbar}{2}\sigma_1, \quad \frac{\hbar}{2}\sigma_2, \quad \frac{\hbar}{2}\sigma_3$$

4.10.1 This is observed in the magnetic splitting of the spectral line of many atoms (e.g., Sodium). Of course, in classical mechanics, it would be impossible for a particle at rest to carry angular momentum.

4.11 The state of an electron is described by functions $\psi : R^3 \to C^2$; the total angular momentum is the sum of the orbital and spin angular momenta:

$$\boldsymbol{J} = \boldsymbol{r} \times \boldsymbol{p} + \frac{\hbar}{2}\boldsymbol{\sigma}$$

4.11.1 Recall that $\boldsymbol{p} = -i\hbar\frac{\partial}{\partial \boldsymbol{r}}$

4.12 There is a one-one correspondence between traceless hermitean matrices and vectors

$$\hat{u} = \begin{pmatrix} u_3 & u_1 + iu_2 \\ u_1 - iu_2 & -u_3 \end{pmatrix} = u_i\sigma_i.$$

4.12.1 The determinant of the matrix is the negative of the square of the length of the corresponding vector:

$$\det \hat{u} = -(u_1^2 + u_2^2 + u_3^2) = -(u, u).$$

4.13 There is a $2 \to 1$ onto homomorphism $SU(2) \to SO(3)$:

$$g\hat{u}g^{-1} = \widehat{R(g)u}$$

4.13.1 In other words,

$$g\sigma_i g^{-1} = R_{ji}(g)\sigma_j$$

4.13.2 A homomorphism between groups is a map that preserves the multiplication law and identity.

4.13.3 Proof that the multiplication is preserved:

$$R(g_1g_2)_{kj}\sigma_k = (g_1g_2)\sigma_i(g_1g_2)^{-1} = R_{ji}(g_2)g_1\sigma_j g_1^{-1} = R_{ji}(g_2)R_{kj}(g_1)\sigma_k = [R(g_1)R(g_2)]_{ki}\sigma_k$$

4.13.4 Proof that the matrix $R(g)$ is orthogonal:

$$\det[g\hat{u}g^{-1}] = \det\hat{u} \Rightarrow (R(g)u, R(g)u) = (u,u).$$

4.13.5 The matrix $R(g)$ is of determinant one because it is a continuous function of g and $R(1) = 1$ has determinant one. (Recall that $\det R(g) = \pm 1$.

4.13.6 Any hermitean matrix can be diagonalized. Thus there is a rotation $R(g)$ such that any vector can be brought to the form $\begin{pmatrix} u_3 & 0 \\ 0 & -u_3 \end{pmatrix}$. Thus the map is onto: any rotation can be realized as $R(g)$ for some g

4.13.7 Both g and $-g$ go to the same matrix $R(g)$: the map is $2 \to 1$.

4.14 A *representation* of a group is a homomorphism to a group of matrices.

4.14.1 Thus the above map $R(g)$ is a representation of $SU(2)$ on three dimensional Euclidean space.

4.15 A *spinor* is a function $\psi : R^3 \to C^2$. The space of spinors carry a representation of $SU(2)$:

$$r(g)\psi(x) = g\psi(R(g)^{-1}x).$$

This is also a projective representation of the rotation group.

4.15.1 A *projective representation* is a map into transformations of a projective space;that is, they act on rays of a vector space rather than vectors.

Chapter 5

Lorentz Invariance

5.1 It is an astonishing physical fact that the speed of light (in vacuum) is the same in all reference frames.

5.1.1 This is not true of other waves; for example the speed of sound is different for observers moving at different velocities with respect to air.

5.2 The law of addition of velocities has to be modified to live be consistent with this peculiar result:

$$v_1 \oplus v_2 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}.$$

5.2.1 If we require that the addition of velocities forms a group, and that c added to any velocity give c , this is the only answer.

5.2.2 The easiest way to check the group property to make the change of variable

$$\frac{v}{c} = \tanh\theta$$

and use the addition formula for $\tanh$.

5.3 This means that the shape of the wavefront of light is the same for all observers.

5.3.1 If a pulse of light is emitted at the origin at time $t = 0$, it will spread along the cone

$$c^2t^2 - x_1^2 - x_2^2 - x_3^2 = 0, t > 0.$$

5.4 The laws of physics have to the same for all observers moving at constant velocity relative to each other.

5.4.1 This innocuous statement has important consequences when combined with the fact the the velocity of light is the same for all such observers.

5.4.2 The Newtonian notions of time, space, energy, momentum all need to be modified.

5.4.3 Minkowski noted that the addition law for velocities has a simple geometric interpretation: the rule for the distance between two points in space-time is $\sqrt{c^2(t-t')^2 - (x_1 - x_1')^2 - (x_2 - x_2')^2 - (x_3 - x_3')^2}$

5.4.4 If velocity is identified as $\frac{dx}{dt}$, the rule for addition of velocities corresponds to 'rotations' around an imaginary angle. A more precise version is,

5.5 The *Minkowski* inner product of a pair of vectors in space-time is

$$(u, v) = u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3.$$

5.5.1 We can also write this in matrix notation:

$$(u, v) = u^T\eta v \quad \eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

5.6 A *Lorentz transformation* is a linear transformation leaving this inner product invariant:

$$[\Lambda u]^T \eta [\Lambda v] = u^T \eta v, \quad \Rightarrow \Lambda^T \eta \Lambda = \eta.$$

5.6.1 If Λ_1 and Λ_2 satisfy this equation, so will the product and inverse: the set of Lorentz transformations is a group.

5.6.2 $SO(3)$ is a subgroup:

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}, \quad R^T R = 1.$$

5.6.3 By taking determinants it is easy to see that the determinant of a Lorentz matrix is either ± 1 . The sign of the determinant is constant under continuous changes of matrix elements.

5.6.4 The 00 component of the condition says

$$\Lambda_{00}^2 - \Lambda_{01}^2 - \Lambda_{02}^2 - \Lambda_{03}^2 = 1.$$

This is the equation for a hyperboloid of two sheets. So the sign of Λ_{00} is constant under continuous changes of the matrix elements

5.6.5 The set of Lorentz transformations splits into four connected components labelled by the signs of $\det \Lambda$ and Λ_{00} .

5.6.6 The component containing the identity (i.e., $\det \Lambda = 1$ and $\Lambda_{00} > 0$) is the group of *proper Lorentz transformations*. They preserve the orientation of time and do not contain reflection of an odd number of spatial directions.

5.7 Only the proper Lorentz transformations are symmetries of physical laws.

5.7.1 Weak interactions are not invariant under violate parity as well as time reversal.

5.8 The relation of momentum $\mathbf{p}$ to energy E is

$$(p,p) = m^2c^4, \quad p = (E, c\mathbf{p}).$$

5.8.1 When $m > 0$ this is a two-sheeted hyperboloid; if $m = 0$ this is a cone.

5.8.2 The case $m^2 < 0$ is un-physical since the sign of energy is not invariant under proper Lorentz transformations. The hyperboloid in this case has a single sheet and contains energies that are negative: some observers will see that the energy of a given vector is is positive and others will see it as negative, so there is no way to exclude negative energies. But then we can make energies as negative as we want, and the system is unstable against emission of arbitrarily large amounts of energy. These non-existent particles are called *tachyons*.

Chapter 6

Bohr-Sommerfeld Quantization

6.1 Sommerfeld generalized the Bohr model of the atom to include relativistic effects. He found that the energy levels predicted agree even better: they explain the hyperfine structure.

6.1.1 In the Bohr model the energies only depend on the principal quantum number, which is the sum of the radial and angular momentum quantum numbers. Relativistic effects remove this degeneracy, so that the energy does depend separately on the principal and angular momentum quantum numbers. This dependence was already known in atomic spectroscopy as a small effect of about 0.01% . This was Sommerfeld's achievement. The enery levels are still degenerate: they remain independent of the magnetic quantum number, a consequence of rotaion invariance.

6.1.2 The mass shell condition becomes

$$(p_0 - V)^2 - c^2 p^2 = m^2 c^4, \quad V = \frac{Ze^2}{r}$$

since $p_0 - V$ is the kinetic energy. In radial co-ordinates

$$p^2 = p_r^2 + \frac{p_\theta^2}{r^2}$$

where p_θ is the angular momentum and p_r the radial momentum. Since p_θ is conserved we treat it as a constant. The Bohr condition says that $\int p_\theta d\theta = 2\pi l\hbar$ for some integer l .

6.1.3 The mass-momentum relation yields, for $V = -\frac{Ze^2}{r}$

$$[p_0 - V]^2 - c^2\left[p_r^2 + \frac{p_\theta^2}{r^2}\right] = m^2c^2 \quad \Rightarrow$$

$$p_r = \left[\frac{1}{c^2}(p_0 - V)^2 - m^2c^2 - \frac{p_\theta^2}{r^2}\right]^{\frac{1}{2}} = \left[-A + \frac{B}{r} - \frac{C}{r^2}\right]^{\frac{1}{2}}$$

where

$$A = m^2c^2 - \frac{p_0^2}{c^2}, \quad B = \frac{2Ze^2p_0}{c^2}, \quad C = p_\theta^2 - \frac{Z^2e^4}{c^2}.$$

6.1.4 The Bohr-Sommerfeld quantization condition says that the total action of a complete orbit is an integer multiple of Plank's constant $h = 2\pi\hbar$:

$$2\int_{r_1}^{r_2} p_r dr = 2\pi n_r \hbar$$

where r_1 and r_2 are the turning points at which p_r vanishes.

6.1.5 The integral can be evaluated by countour integral methods:

$$2\int_{r_1}^{r_2}\left(-A + \frac{B}{r} - \frac{C}{r^2}\right)^{\frac{1}{2}} dr = -2\pi\left[\sqrt{C} + \frac{B}{2\sqrt{A}}\right]$$

6.1.6 Thus the energy p_0 is the solution of

$$n_r\hbar = -\left[\left(p_\theta^2 - \frac{Z^2e^4}{c^2}\right)^{\frac{1}{2}} + \frac{1}{2}\frac{2Ze^2p_0}{c^2}\left(m^2c^2 - \frac{p_0^2}{c^2}\right)^{-\frac{1}{2}};\right]$$

i.e.,

$$p_0 = \frac{mc^2}{\sqrt{1+a}}, \quad a = \frac{Z^2e^4}{c^3}\frac{1}{\left(\hbar n_r + \sqrt{\hbar^2 l^2 - \frac{Z^2e^4}{c^2}}\right)^2}$$

In the limit of large c this reduces to the usual formula for hydrogen.

Chapter 7

The Klein-Gordon Equation

7.1 The set of four-momenta of a particle with mass m form a hyperboloid:

$$p_0^2 - c^2p_1^2 - c^2p_2^2 - c^2p_3^2 = m^2c^4, \quad p_0 > 0.$$

7.2 If we ignore the condition that the energy has to be positive, we can express this as a simple differential equation for its wave-function:

$$-\hbar^2[\frac{\partial^2}{\partial t^2} - c^2\frac{\partial^2}{\partial x_1^2} - c^2\frac{\partial^2}{\partial x_2^2} - c^2\frac{\partial^2}{\partial x_3^2}]\psi = m^2c^4\psi$$

7.2.1 Recall that $p_0 = i\hbar\frac{\partial}{\partial t}, p_1 = -i\hbar\frac{\partial}{\partial x^1}$ etc. in quantum mechanics.

7.2.2 The equation is invariant under Lorentz transformations-including parity and time reversal.

7.2.3 The equation allows for negative energy solutions. To properly interpret this situation, we need quantum field theory. We will return to this topic later.

7.2.4 Under the influence of a potential V the mass-shell condition changes to

$$(p_0 - V)^2 - c^2p^2 = m^2c^4.$$

7.2.5 The wave equation of relativistic quantum mechanics corresponding to this is

$$\left[i\hbar\frac{\partial}{\partial t}-V\right]^2\psi+c^2\hbar^2\nabla^2\psi=m^2c^4\psi.$$

7.2.6 A negatively charged pi meson can form a bound state with a nucleus analogous to the hydrogen atom. Althuogh relativistic effects are small, they have been observed in classic experiments of Wu et al in the late 1970's. We can determine the 'fine structure' of pionic atoms by solving the Klein-Gordon equation in a Coulomb potential $V(r)=-\frac{Ze^2}{r}$.

7.2.7

$$\left[i\hbar\frac{\partial}{\partial t}-V(r)\right]^2\psi+c^2\hbar^2\left[\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\nabla_\theta^2\right]\psi=m^2c^4\psi.$$

where ∇_θ^2 is the usual angular part of the Laplace operator.

7.2.8 Recall that the spherical harmonics $Y_{lm}(\theta,\phi)=P_l^{|m|}(\theta)e^{im\phi}$ are the eigenfunctions of the angular Laplacian:

$$\nabla_\theta^2 Y_{lm}(\theta,\phi)=-l(l+1)Y_{lm}(\theta,\phi)$$

7.2.9 As in the non-relativistic theory, we can solve this equation by postulating a separation of variables

$$\psi(r,\theta,\phi)=rR(r)Y_{lm}(\theta,\phi)e^{-\frac{iEt}{\hbar}}$$

to get the differential equation for the radial function

$$R''-\left[\frac{m^2c^4-(E-V)^2}{\hbar^2c^2}+\frac{l(l+1)}{r^2}\right]R=0$$

7.2.10 For the Coulomb potential $V(r) = -\frac{Ze^2}{r}$ we get

$$R'' + \left[-A + \frac{B}{r} - \frac{C}{r^2}\right] R = 0$$

with

$$A = \frac{m^2c^4 - E^2}{\hbar^2 c^2}, \quad B = \frac{2EZe^2}{\hbar^2 c^2}, \quad C = l(l+1) - \frac{Z^2 e^4}{\hbar^2 c^2}.$$

7.2.11 Fortunately this is the *same* equation we would get in the non-relativistic case, but with different constants.There we would have

$$A = \frac{2m|E|}{\hbar^2}, \quad B = \frac{2mZe^2}{\hbar^2}, \quad C = l(l+1).$$

So the same methods will apply to our relativistic case.

7.2.12 The constant $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$ is a dimensionless combination of fundamental constants of nature, called the *fine structure constant*. It determines the size of the relativistic corrections, the 'fine structure' of atomic spectral lines.

7.2.13 It is sometimes convenient to use so-called 'natural units' in which $\hbar = c = 1$. Then, energy has dimensions of inverse length, a does mass and momentum. In these units,

$$A = m^2 - E^2, \quad B = 2EZ\alpha, \quad C = l(l+1) - Z^2\alpha^2$$

which helps with the book-keeping.

7.2.14 For large r , $R \sim e^{-r\sqrt{A}}$ so it is useful to introduce the dimensionless variable $\rho = r\sqrt{A}$ and the dimensionaless function of energy

$$\epsilon = \frac{B}{\sqrt{A}} = 2\frac{E}{\sqrt{m^2 - E^2}} Z\alpha$$

to get

$$\frac{d^2 R}{d\rho^2} + \left[-1 + \frac{\epsilon}{\rho} - \frac{C}{\rho^2}\right] R = 0.$$

7.2.15 For small r , $R(r) \sim r^{\lambda+1}$ with

$$\lambda(\lambda+1) = C = l(l+1) - Z^2\alpha^2.$$

Note that because of the relativistic correction to C , the exponent λ need not be an integer;since $Z\alpha << 1$, this shift away from an integer is small. We discard the other solution $\rho^{-\lambda}$ as it blows up at the origin.

7.2.16 Combining all the above, we make the change of variable $R\left(\frac{\rho}{\sqrt{A}}\right) = \rho^{\lambda+1}e^{-\rho}w(\rho)$ to get

$$w'' + 2\left[\frac{\lambda+1}{\rho} - 1\right]w' + \frac{\epsilon - 2(\lambda+1)}{\rho}w = 0.$$

A power series $w(\rho) = \sum_{k=0}^{\infty} a_k\rho^k$ gives the recusion

$$\frac{a_{k+1}}{a_k} = \frac{-\epsilon + 2(k+\lambda+1)}{(k+\lambda+1)(k+\lambda+1) - \lambda(\lambda+1)}$$

since this ratio goes like $\frac{1}{k}$for large k the series will not converge; unless it truncates at some value of k so that w is really a polynomial. The condition for this is

$$\epsilon = 2(k+\lambda+1).$$

which determines the energy. The rest is algebra.

7.2.17 Te algebraic relation that determines the spectrum is

$$\frac{B}{\sqrt{A}} = 2(k+\lambda+1), \quad \lambda(\lambda+1) = C$$

or

$$A = \frac{B^2}{4(k+\lambda+1)^2}, \quad k = 0, 1, \cdots.$$

This quadratic equation can be solved for energy.

7.2.18 The positive energy solutions are in agreement with observed values for pionic atoms; the discrepancies that remain can be explained from Quantum Electrodynamics.

7.2.19 There are also negative energy solutions. These cause many conceptual problems as they don't exist in nature. The situation is even worse. They cause an instability of the atom where the electron drops into the negative energy states releasing arbitrarily large amounts of energy. The resolution to this lies in Quantum Field Theory, which totally revamps the foundations of physics.

Chapter 8

Spinors in Space-Time

8.1 The group $SL(2,C)$ is the set of complex 2×2 matrices of determinant one.

8.1.1 A subgroup is the group $SU(2)$ of special unitary matrices.

8.1.2 Three complex numbers or six real numbers are needed to determine an element of $SL(2,C)$.

8.2 There is a $2\to 1$ homomorphism from $SL(2,C)$ to the proper Lorentz group.

8.2.1 We begin with the observation that

$$p^T\eta p = \det\hat p,\quad \hat p = \begin{pmatrix} p_0-p_3 & -p_1-ip_2 \\ -p_1+ip_2 & p_0+p_3\end{pmatrix} = p_0 1 - \sigma_1 p_1 - p_2\sigma_2 - p_3\sigma_3.$$

Moreover, from any hermitean matrix we can extract such a vector uniquely.

8.2.2 Now for each 2×2 matrix g , there is a 4×4 matrix $\Lambda(g)$ such that

$$g\hat p g^\dagger = \widehat{\Lambda(g)}p.$$

since the l.h.s. is also a hermitean matrix which depends linearly on p . If moreover $\det g = 1$, we will have $\Lambda(g)^T\eta\Lambda(g) = \eta$, since

$$\det g\hat p g^\dagger = \det\hat p, \Rightarrow [\Lambda(g)p]^T\eta[\Lambda(g)p] = p^T\eta p.$$

8.2.3 Since $SL(2,C)$ is connected and Λ is continuous, the image has to be contained in the connected component of the identity in the Lorentz group. Since any hermitean matrix can be diagonalized, any four-vector can be reduced to the form $(p_0, p_3, 0, 0)$ by some transformation $\Lambda(g)$ in the image of this map. This can be seen to imply that the image is the whole connected component of the identity in the Lorentz group.

8.2.4 Both g and $-g$ are mapped to the same element in the Lorentz group. The kernel of the map is exactly $\{1, -1\}$.

8.3 A map from a group to a set of matrices on a vector space is called a *representation* if it preserves the multiplication law. Two representations are said to be *equivalent* if they differ only by an equivalence transformation; i.e., if there is a matrix S such that

$$r_1(g) = Sr_2(g)S^{-1}.$$

8.3.1 Thus we have a representation of $SL(2,C)$ on four-vectors of space-time.

8.3.2 The defining representation of $SL(2,C)$ is simply the map g to itself. The map $g \mapsto g^{\dagger -1}$ is also a representation.

8.3.3 This is not equivalent to the defining representation; there can be no matrix satisfying $g^{\dagger -1} = SgS^{-1}$ because there are cases where g and $g^{\dagger -1}$ have different eigenvalues. (For example $g = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ with a complex non-zero a .)

8.3.4 The map $g \mapsto g^*$ is also a representation. But for two by two matrices

$$g^{\dagger -1} = \sigma_2 g^* \sigma_2^{-1}$$

so this equivalent to the representation mentioned above!

8.3.5 Thus we see that $SL(2,C)$ has two different (inequivalent) representations on C^2 . These are called the spinor representations; the defining representation is called the right handed spinor and its conjuagte is the left-handed spinor. (Beware! conventions vary on this.)

8.3.6 Multiplication by $\hat{p}$ turns a left handed spinor to a right handed spinor:

$$\hat{p} \mapsto \widehat{\Lambda p} = g\hat{p}g^{\dagger}, \quad \psi \mapsto g^{\dagger -1}\psi, \quad \hat{p}\psi \mapsto g\hat{p}\psi.$$

8.4 The *Pauli equation*

$$\hat{p}\psi(p) = 0$$

is invariant under the action of the Lorentz group

$$\psi(p) \to g^{\dagger}\psi(\Lambda(g^{-1})p)$$

8.4.1 This equation is not invariant under parity.

8.5 Th Pauli equation describes particles that are massless and of spin $\frac{1}{2}$: a good approximation for neutrinos.

8.6 The Pauli equation is a differential equation when the usual rules $p_\mu = i\hbar\frac{\partial}{\partial x^\mu}$ are applied:

$$\begin{pmatrix} \partial_0 + \partial_3 & \partial_1 + i\partial_2 \\ \partial_1 - i\partial_2 & \partial_0 - \partial_3 \end{pmatrix} \psi = 0$$

8.6.1 By applying the conjugate operator $\begin{pmatrix} \partial_0 - \partial_3 & -\partial_1 - i\partial_2 \\ -\partial_1 + i\partial_2 & \partial_0 + \partial_3 \end{pmatrix}$ on the left we see that each component of the spinor satisfies the wave equation: the signals described by ψ propagate at the speed of light.

8.6.2 Its hermitean conjugate satisfies

$$\partial_\mu \psi^{\dagger}\sigma^{\mu} = 0.$$

8.6.3 The quantity $j_\mu = \psi^\dagger \sigma_\mu \psi$ transforms as a vector field. The Pauli equation implies that this vector field is conserved:

$$\eta^{\mu\nu}\partial_\mu j_\nu = 0, \quad \partial_0 j_0 - \partial_1 j_1 - \partial_2 j_2 - \partial_3 j_3 = 0.$$

8.6.4 If we integrate this over all of space at a given instant in time

$$\frac{\partial}{\partial x^0}\int j_0 d^3x = \int \nabla \cdot j d^3x = 0$$

by Gauss' theorem. (We assume that ψ vanishes at infinity so that the integral on the l.h.s., $\int \psi^\dagger \psi d^3x$ is finite.) Thus $\int \psi^\dagger \psi d^3x$ is independent of time. This density $\psi^\dagger \psi$ can be thought of as the probability density of finding the particle described by ψ .

8.6.5 In spite of being first order in time, the Pauli equation has negative energy solutions. This requires that we change the entire physical basis by passing to quantum field theory.

8.6.6 On passing to the quantum field theory, the energy of the particles is positive while the density $\psi^\dagger \psi$ becomes non-positive: it describes the net number of particles minus anti-particles.

8.6.7 It is possible to modify the Pauli equation to allow for a non-zero mass, but then we lose this conserved quantity.

8.7 A Lorentz invariant wave equation for a massive particle of spin $\frac{1}{2}$ is

$$-i\sigma \cdot \partial\psi = m\sigma_2\psi^*.$$

This is often called the Majorana equation.

8.7.1 Recall that $\psi \mapsto g^{\dagger -1}\psi$ and $\sigma \cdot p\psi \mapsto g\sigma \cdot p\psi$ under a Lorentz transformation; that is, $\psi \mapsto \sigma_2 g^* \sigma_2 \psi$.

8.7.2 Thus

$$\psi^* \mapsto \sigma_2 g \sigma_2 \psi^*, \quad \sigma_2\psi^* \mapsto g\sigma_2\psi^*$$

and the l.h.s. and r.h.s. transform the same way under Lorentz transformations.

8.7.3 These describe massive spin half particles that are there own anti-particles: one of the candidates to describe a massive neutrino. The true story about neutrinos is not known yet, since there are several different kinds of them, and they get mixed with each other.

8.7.4 Because it involves both ψ and ψ^* , it is often thought of an equation for the four real components of ψ (the real and imaginary parts of the two complex components) and then the Pauli matrices are turned into four by four real matrices.

8.7.5 Recall that the multiplication of a complex number $a+ib$ by is the same as multiplication of the real vector $\begin{pmatrix} a \\ b \end{pmatrix}$ by the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Thus we can think of the Pauli matrices as for by four real matrices

$$\sigma_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

instead of two by two complex matrices. In this language complex conjugation $\psi \to \psi^*$ would be a real linear transformation by the matrix

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

8.7.6 Exercise: Rewrite the Majorana equation as a linear differential equation for a four component real function on space-time.

Chapter 9

The Dirac Equation

9.1 The Pauli matrices satisfy

$$[\sigma_i, \sigma_j]_+ = 2\delta_{ij}$$

where $[A, B]_+ = AB + BA$ **denotes the anti-commutator.**

9.1.1 All their properties that are independent of the choice of basis can be obtained just from these relations.

9.1.2 For example,

$$\mathrm{tr}\ \sigma_i\sigma_j = \mathrm{tr}\ \sigma_j\sigma_i = \frac{1}{2}\,\mathrm{tr}\ [\sigma_i\sigma_j + \sigma_j\sigma_i] = \delta_{ij}.$$

9.1.3 Every Pauli matrix has a matrix has another that anti-commutes with it; hence its trace is zero; e.g.,

$$\mathrm{tr}\ \sigma_1 = \mathrm{tr}\ \sigma_2^2\sigma_1 = \mathrm{tr}\ \sigma_2\sigma_1\sigma_2 = -\ \mathrm{tr}\ \sigma_1\sigma_2^2 = -\ \mathrm{tr}\ \sigma_1.$$

9.1.4 Any 2×2 matrix is a linear combination of Pauli matrices

$$A = a_0 1 + a_i\sigma_i, \quad a_0 = \frac{1}{2}\,\mathrm{tr}\ A, \quad a_i = \frac{1}{2}\,\mathrm{tr}\ A\sigma_i.$$

Note that the number of independent matrix elements of A is equal to the number of components a_0, a_i .

9.1.5 In particular $\sigma_1\sigma_2 = \sigma_3$ etc. can be deduced.

9.1.6 A two by two matrix that commutes with every Pauli matrix is a multiple of the identity.

9.2 The *Dirac matrices* are defined by

$$[\gamma_\mu, \gamma_\nu]_+ = 2\eta_{\mu\nu}, \quad \mu, \nu = 0, 1, 2, 3.$$

9.2.1 The analogy with Pauli matrices is clear.

9.2.2 In detail, $\gamma_0^2 = 1, \gamma_0\gamma_i = -\gamma_i\gamma_0, \gamma_i\gamma_j + \gamma_j\gamma_i = -2\delta_{ij}$.

9.2.3 The eigenvalues of γ_0 are ± 1 . Moreover, $\gamma_0\gamma_1 = -\gamma_1\gamma_0, \gamma_1^2 = -1 \Rightarrow \gamma_0 = -\gamma_1^{-1}\gamma_0\gamma_1 \Rightarrow \operatorname{tr}\,\gamma_0 = 0.$ So the degeneracy of ± 1 are equal (say n). There is a basis in which

$$\gamma_0 = \begin{pmatrix} 1_n & 0_n \\ 0_n & -1_n \end{pmatrix}.$$

9.2.4 A matrix that anti-commutes with $\begin{pmatrix} 1_n & 0_n \\ 0_n & -1_n \end{pmatrix}$ is of the form $\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}$. It follows that

$$\gamma_0 = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix}, \quad \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}$$

satisfy the Dirac anti-commutation relations.

9.3 If γ_μ satisfy the Dirac relations, so will $S\gamma_\mu S^{-1}$ for any invertible S .

9.3.1 Proof is obvious. This corresponds to a change of basis.

9.4 The converse is true as well: any set of 4×4 matrices γ_μ , γ'_μ satisfying the Dirac relations are connected by a change of basis $\gamma'_\mu = S\gamma_\mu S^{-1}$.

9.4.1 Any such matrix can be expanded in terms of the Dirac matrices and their products: there are 2^4 independent components.

9.5 Any product of Dirac matrices can be written as a linear combination of the 16 matrices

$$1, \gamma_\mu, \gamma_\mu\gamma_\nu \text{ for } \mu < \nu, \gamma_\mu\gamma_\nu\gamma_\rho \text{ for } \mu < \nu < \rho, \quad \gamma_0\gamma_1\gamma_2\gamma_3.$$

9.5.1 It follows that the minimum dimension of matrices satisfying the Dirac algebra is 4 .

9.5.2 Each of these (except 1) has zero trace. Any product of two distinct matrices above has zero trace;

9.5.3 Any 4×4 matrix can be expanded in terms of the Dirac matrices and their products: there are 16 independent components.

9.5.4 Define

$$\sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]$$

9.5.5 These satisfy the commutation relations of the infinitesimal Lorentz transformations:

$$[\frac{1}{2}\sigma_{\mu\nu}, \frac{1}{2}\sigma_{\rho\sigma}] = \eta_{\nu\rho}\frac{1}{2}\sigma_{\mu\sigma}$$

Moreover, the Dirac matrices transform as vectors under it.

$$[\frac{1}{2}\sigma_{\mu\nu}, \gamma_\sigma] = \eta_{\nu\sigma}\gamma_\mu - \eta_{\mu\sigma}\gamma_\nu$$

9.5.6 **Hint**: It is actually easier to start by establishing the second relation and then use it to establish the first.

9.6 The Dirac equation

$$[i\gamma\cdot\partial+m]\psi=0$$

is invariant under the infinitesimal Lorentz transformations generated by

$$J_{\mu\nu}=[x_\mu\partial_\nu-x_\nu\partial_\mu]+\frac{1}{2}\sigma_{\mu\nu}.$$

9.6.1 We verify that under $J_{\mu\nu}$ both ∂_μ and γ_μ transform as Lorentz vectors so that the scalar product is invariant:

$$[J_{\mu\nu},\partial_\sigma]=\eta_{\nu\sigma}\partial_\mu-\eta_{\mu\sigma}\partial_\nu$$

$$[J_{\mu\nu},\gamma_\sigma]=\eta_{\nu\sigma}\gamma_\mu-\eta_{\mu\sigma}\gamma_\nu$$

$$[J_{\mu\nu},\gamma\cdot\partial]=0.$$

9.7 The Dirac equation describes a particle of mass m .

9.7.1 We use units in which $\hbar=c=1$. The Dirac equation has a plane wave solution

$$\psi(x)=ue^{ip\cdot x}$$

for a constant u if

$$\gamma\cdot pu=mu.$$

This eigenvalue problem has a solution only if $p\cdot p=m^2$; for,

$$[\gamma\cdot p]^2=\frac{1}{2}[\gamma_\mu\gamma_\nu+\gamma_\nu\gamma_\mu]p^\mu p^\nu=p\cdot p.$$

9.7.2 There are solutions with positive as well as negative energy. Indeed, if $p = (p_0, \mathbf{0})$,

$$p_0 \gamma_0 u = mu \Rightarrow p_0 = m, u = \begin{pmatrix} u_1 \\ u_2 \\ 0 \\ 0 \end{pmatrix}, \text{or } p_0 = -m, u = \begin{pmatrix} 0 \\ 0 \\ u_3 \\ u_4 \end{pmatrix}.$$

The meaning of the negative energy solutions will be made clear later through the hole theory of Dirac.

9.8 The particle described by the Dirac equation has spin $\frac{1}{2}$.

9.8.1 We look at the generators of angular momentum contained in the Lorentz transformation above:

$$J_{ij} = [x_i \partial_j - x_j \partial_i] + \frac{1}{2} \sigma_{ij}$$

Using the explicit representation of the Dirac matrices and identifying $J_1 = -iJ_{12}$ (so that the angular momentum is represented by hermitean operators) we get

$$\mathbf{J} = \mathbf{L} + \frac{1}{2} \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$$

and $\mathbf{L}$ is the orbital angular momentum operator of quantum mechanics.

Chapter 10

The Hole Theory of Dirac

10.0.2 Dirac proposed a dramatic reinterpretation of physics to avoid the contradictions inherent in the negative energy solutions.

10.1 Particles of spin $\frac{1}{2}$ obey the exclusion principle: only one such particle can occupy a state.

10.1.1 We know that the electrons obey the exclusion principle from atomic physics: the shell model of atomic orbitals explains the periodic table of elements.

10.2 Axiom: The 'vacuum' state (the state of the lowest possible energy) already has all the negative energy states occupied.

10.2.1 It will cost energy to either create a positive energy particle or remove a negative energy particle.

10.3 Axiom: The electric charge of the electrons occupying the negative energy states is not measurable.

10.3.1 The absence of a negative energy particle (a *hole*) itself behaves like a new particle of positive energy and with the opposite electric charge.

10.3.2 A particle-hole pair can be produced if we have at least $2m$ units of energy available.

10.3.3 A constant electric field E extending over a distance L will be unstable due to production of particle-hole pairs if $EL > 2m$. But since $m \sim .511$ MeV for the electron, this is far beyond the capacity of materials. Such fields might exist in outer space and could be the source of some of the cosmic rays.

10.3.4 A photon of energy greater than $2m$ (in the laboratory frame) can produce a particle-hole pair if it collides with a nucleus. This was observed, originally using cosmic gamma rays .

Chapter 11

The Dirac Equation in a Centrally Symmetric Potential

11.1 The exact determination of the energy levels of hydrogen is the foundation of atomic physics.

11.1.1 At first we solve the non-relativistic Schrödinger equation in the Coulomb field of the nucleus. The relativistic effects are small because the velocity of the electron in the hydrogen atom is $\alpha \sim \frac{1}{137}$ .So the change in energy levels due to relativistic effects will be down by a factor of $\sim 10^{-4}$.

11.1.2 Although small the relativistic effects of hydrogen are easily visible in spectroscopy: the fine structure. The energies are independent of the angular momentum in non-relativistic quantum mechanics; the small observed dependence on the angular momentum was explained by Sommerfeld as by a relativistic generalization of the Bohr model.

11.2 We will need to solve the Dirac equation in a spherical symmetric potential to understand the fine structure of a hydrogenic atom.

11.3 The key idea is to exploit the conservation of angular momentum to reduce the problem to one dimension.

11.3.1 We will see that only the sum of orbital and spin angular momentum is conserved.

11.4 The Dirac equation of a particle in a static potential V reduces to

$$[\alpha \cdot \mathbf{p} + \beta m + V]\psi = E\psi.$$

where

$$\alpha_i = \gamma_0\gamma_i, \quad \beta = \gamma_0, \mathbf{p} = -i\nabla$$

11.4.1 The free Dirac equation can be written as

$$[i\gamma_0\partial_0 - \gamma \cdot \mathbf{p} + m]\psi = 0$$

For a stationary solution $\psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEx^0}$,

$$E\psi = [\gamma_0\gamma \cdot \mathbf{p} + \gamma_0 m]\psi$$

Thus the operator $\alpha \cdot \mathbf{p} + m\beta$ represents the Kinetic energy. In a potential V , we must put this equal to $E - V$.

11.4.2 These matrices satisfy

$$[\alpha_i, \alpha_j]_+ = 2\delta_{ij}, \quad \beta^2 = 1, \quad [\beta, \alpha_i]_+ = 0.$$

11.4.3 An explicit representation would be

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}, \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

11.5 If the potential V depends only the distance from the origin, and is spin independent, the total angular momentum

$$\mathbf{J} = \mathbf{r} \times \mathbf{p} + \frac{1}{2}\Sigma$$

commutes with the Dirac hamiltonian $H = \alpha \cdot \mathbf{p} + \beta m + V$:

$$[\mathbf{J}, H] = 0.$$

11.5.1 Here, Σ are the spin matrices satisfying

$$[\frac{1}{2}\Sigma_i, \alpha_j] = \sqrt{-1}\epsilon_{ijk}\alpha_k, \quad [\Sigma_i, \beta] = 0.$$

Explicitly

$$\alpha_i\alpha_j = \delta_{ij} + \frac{1}{2}\sqrt{-1}\epsilon_{ijk}\Sigma_k$$

$$\Sigma_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}.$$

11.5.2 Spin independence means that V is proportional to the identity in the space of spinors. It is clear that a spherically symmetric potential will satisfy

$$[\mathbf{L}, V] = 0, [\Sigma, V] = 0$$

separately so that $[\mathbf{J}, V] = 0$.

11.5.3 Since α and $\mathbf{p}$ transform as vectors under $\mathbf{J}$ it follows that $\alpha \cdot \mathbf{p}$ is a scalar: $[\mathbf{J}, \alpha \cdot \mathbf{p}] = 0$. Similarly $[\mathbf{J}, \beta] = 0$. Piecing all this together we get

$$[\mathbf{J}, H] = 0.$$

Note that neither L_i nor Σ_i commutes with $\alpha \cdot \mathbf{p}$, hence with H .

11.5.4 We now separate the Dirac operator into parts involving radial and angular derivatives. Let us start with the identity

$$\mathbf{p} = \mathbf{r}\ \frac{\mathbf{r}}{r^2} \cdot \mathbf{p} - \frac{1}{r^2}\mathbf{r} \times (\mathbf{r} \times \mathbf{p})$$

The cross-product of $\mathbf{p}$ with the unit vector $\frac{\mathbf{r}}{r}$ rotates its component orthogonal to $\mathbf{r}$ through a right angle; another such cross product will give the negative of this component, explaining the sign of the second term.

11.5.5 Thus

$$\alpha \cdot \mathbf{p} = \alpha_r p_r - \frac{1}{r^2}(\alpha \times \mathbf{r}) \cdot \mathbf{L}$$

where $\alpha_r = \frac{1}{r}\mathbf{r} \cdot \alpha$ etc.

11.5.6 It is easy to see that $\alpha_r^2 = 1$. More generally, $\alpha_i \alpha_j = \delta_{ij} + i\epsilon_{ijk}\Sigma_k$ gives

$$\alpha_r \boldsymbol{\alpha} = \frac{\mathbf{r}}{r} - i\frac{\mathbf{r}}{r} \times \boldsymbol{\Sigma}$$

11.5.7 Using $(\boldsymbol{r} \times \boldsymbol{\Sigma}) \times \boldsymbol{r} = r^2\boldsymbol{\Sigma} - \Sigma_r r\boldsymbol{r}$,

$$(\alpha_r \boldsymbol{\alpha} \times \frac{\mathbf{r}}{r}) \cdot \mathbf{L} = -i\boldsymbol{\Sigma} \cdot \mathbf{L}.$$

11.5.8 Hence

$$\boldsymbol{\alpha} \cdot \mathbf{p} = \alpha_r \left(p_r - i\frac{1}{r}\boldsymbol{\Sigma} \cdot \mathbf{L} \right)$$

11.5.9 The Dirac equation becomes

$$\left[-i\alpha_r \left(ip_r + \frac{1}{r}\boldsymbol{\Sigma} \cdot \mathbf{L} \right) + m\beta + (V - E) \right] \psi = 0.$$

Now, in radial co-ordinates the measure of integration is $r^2 dr d\Omega$. An integration by parts will show that $\partial_r^\dagger = -r^{-2}\partial_r r^2 = -\partial_r - \frac{2}{r} \Rightarrow \left(\partial_r + \frac{1}{r}\right)^\dagger = -\left(\partial_r + \frac{1}{r}\right)$. So we must identify $ip_r = \partial_r + \frac{1}{r}$. If we define $K = -[1 + \boldsymbol{\Sigma} \cdot \mathbf{L}]$,

$$\left[-i\alpha_r \left(\partial_r - \frac{1}{r}K \right) + m\beta + (V - E) \right] \psi = 0.$$

11.5.10 We now show that $[\beta, K] = 0$ and $\alpha_r K = -K\alpha_r$. The first statement is obvious. The second is equivalent to $[\alpha_r, \mathbf{\Sigma}\cdot\mathbf{L}]_+ = -2\alpha_r$. This follows from $[\alpha_i, \Sigma_j]_+ \propto 2\delta_{ij}$ and $[L_i, r_j] = \sqrt{-1}\epsilon_{ijk}r_k$. Thus in the basis where $\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,

$$K = \begin{pmatrix} \hat{k} & 0 \\ 0 & -\hat{k} \end{pmatrix}$$

where

$$\hat{k} = -[1 + \boldsymbol{\sigma}\cdot\mathbf{L}]$$

is the Dirac operator on the sphere.

11.5.11 Since also $(-i\alpha_r)^2 = 1, [\beta, -i\alpha_r]_+ = 0$ we have a basis in which $-i\alpha_r = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. If we let $\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}$,

$$\left[\partial_r - \frac{1}{r}\hat{k}\right]\phi + (V - E - m)\,\chi = 0,$$

$$\left[\partial_r + \frac{1}{r}\hat{k}\right]\chi + (E - V - m)\,\phi = 0,$$

11.5.12 We will now assume that $\phi(\mathbf{r}) = f(r)Y(\frac{\mathbf{r}}{r}), \chi(\mathbf{r}) = g(r)\tilde{Y}(\frac{\mathbf{r}}{r})$,separable into a product of radial and angular functions.

11.5.13 We will show in the next chapter that the spectrum of $\hat{k}$ is the set of non-zero integers. Let Y_k be the eigenfunctions of $K = -[1 + \boldsymbol{\sigma}\cdot\mathbf{L}]$ (*spinorial harmonics*, analogous to spherical harmonics):

$$\hat{k}Y_k = kY_k$$

11.5.14 The Dirac equation in a central potential now reduces to a system of first order equations

$$\begin{pmatrix} f' \\ g' \end{pmatrix} = \begin{pmatrix} \frac{k}{r} & m + E - V \\ m - E + V & -\frac{k}{r} \end{pmatrix}\begin{pmatrix} f \\ g \end{pmatrix}.$$

11.6 When $V = -\frac{Z\alpha}{r}$ as for the Coulomb potential, the Dirac equation can be solved in terms of Confluent Hypergeometric functions.

11.6.1 Exercise: Find the normalizable eigenfunctions (bound states) of the Dirac equation in a Coulomb potential. **Hint**: Express the solution in terms of the Confluent Hypergeometric series; show that for a normalizable solution the series must terminate; determine the eigenvalues from this condition.

11.1 The Confluent Hypergeometric Function

11.7 We will study some infinite series of interest in physics.

11.7.1 The geometric series is defined by

$$f(z) = 1 + z + z^2 + z^3 \cdots$$

It converges in the region $|z| < 1$.

11.7.2 It is easy to check that

$$zf(z) = z + z^2 + z^3 + \cdots = f(z) - 1$$

so that

$$1 + z + z^2 + z^3 \cdots = \frac{1}{1-z}.$$

This allows an analytic continuation outside the unit circle. Indeed, $f(z) \sim \frac{1}{z}$ as $|z| \to \infty$.

11.7.3 The differential equation derived below:

$$f'(z) = 1 + 2z^2 + 3z^+ \ldots, \quad zf'(z) = z + 2z^2 + 3z^3 + \cdots$$

$$\Rightarrow (1-z)f'(z) = 1 + z + z^2 + z^3 + \cdots = f(z)$$

can be solved with the initial condition $f(0) = 1$ to get the same answer.

11.7.4 We can get interesting new series by differentiation:

$$(1-z)^{-2} = 1 + 2z + 3z^2 + 4z^3 + \cdots$$

$$2(1-z)^{-3} = 2 + 3\cdot 2z + 4\cdot 3z^2 + 5\cdot 4z^3 \cdots$$

$$3\cdot 2\ (1-z)^{-4} = 3\cdot 2 + 4\cdot 3\cdot 2z + 5\cdot 4\cdot 3z^2 + 6\cdot 5\cdot 4z^3 + \cdots$$

Rearranging these series,

$$(1-z)^{-3} = 1 + 3\ z + 3\cdot 4\frac{z^2}{2!} + 3\cdot 4\cdot 5\frac{z^3}{3!} + 3\cdot 4\cdot 5\cdot 6\frac{z^4}{4!}\cdots$$

$$(1-z)^{-4} = 1 + 4\ z + 4\cdot 5\frac{z^2}{2!} + 4\cdot 5\cdot 6\ \frac{z^3}{3!} + \cdots$$

11.7.5 We begin to see a pattern:

$$(1-z)^{-n} = 1 + n\ z + n(n+1)\frac{z^2}{2!} + n(n+1)(n+2)\frac{z^3}{3!} + n(n+1)(n+2)(n+3)\frac{z^4}{4!}\cdots$$

This can be established by induction on n : Assuming that the formula above is true for powers up to $-n$,using $\frac{d}{dz}\frac{z^r}{r!} = \frac{z^{r-1}}{(r-1)!}$,

$$\begin{aligned}(1-z)^{-(n+1)} &= \frac{1}{n}\frac{d}{dx}(1-z)^{-n}\\ &= 1 + (n+1)z + (n+1)(n+2)\frac{z^2}{2!} + (n+1)(n+2)(n+3)\frac{z^3}{3!}\cdots\end{aligned}$$

which establishes the formula for the power $-(n+1)$.

11.7.6 We could also have obtained this formula by differentiating the geometric series n times and using

$$\frac{d^n}{dz^n}z^r = r(r-1)(r-2)\cdots(r-n+1)z^{r-n}\cdots, \frac{d^n}{dz^n}(1-z)^{-1} = n!(1-z)^{-(n+1)}$$

to get

$$\begin{aligned}
(1-z)^{-(n+1)} &= \frac{1}{n!}\frac{d^n}{dz^n}\sum_{r=0}^{\infty} z^r \\
&= \sum_{r=n}^{\infty} \frac{r(r-1)(r-2)\cdots(r-n+1)}{n!} z^{r-n} \\
&= \sum_{s=0}^{\infty} \frac{(n+s)(n+s-1)\cdots(s+1)}{n!} z^s \\
&= \sum_{s=0}^{\infty} \frac{(n+s)!}{n!s!} z^s \\
&= \sum_{s=0}^{\infty} (n+1)(n+2)\cdots(n+s)\frac{z^s}{s!}
\end{aligned}$$

11.7.7 These manipulations are justified because the geometric series converges inside the unit circle; so it defines a analytic function whose Taylor series can be differentiated term by term.

11.7.8 We have just discovered a version of the binomial theorem.The formula is true even for fractional (indeed complex) values of the exponent:

$$(1-z)^{-a} = 1 + a\ z + a(a+1)\frac{z^2}{2!} + a(a+1)(a+2)\frac{z^3}{3!} + a(a+1)(a+2)(a+3)\frac{z^4}{4!}\cdots$$

Indeed this can be thought of as obtained from the geometric series by differentiating a fractional number of times! The fractional derivative of a power is *defined* to be

$$D^a z^r = r(r-1)(r-2)\cdots(r-a+1)z^{r-a}$$

This will satisfy the usual rules

$$D^a D^b = D^{a+b}$$

11.7.9 It is interesting to take limit as a tends to infinity. To get a sensible limit we must simultaneously replace z by $\frac{z}{a}$. Now recall the 'compound interest' definition of the exponential:

$$\exp(z) = \lim_{a\to\infty}\left(1-\frac{z}{a}\right)^{-a}$$

Note also that

$$a(a+1)(a+2)(a+3)\cdots(a+r-1)\left(\frac{z}{a}\right)^r = \left(1+\frac{1}{a}\right)\left(1+\frac{2}{a}\right)\left(1+\frac{3}{a}\right)\cdots\left(1+\frac{r-1}{a}\right)z^r \to$$

as $a \to \infty$ .Applying this to the series

$$\left(1-\frac{z}{a}\right)^{-a} = \sum_{r=0}^{\infty}\frac{1}{r!}a(a+1)(a+2)(a+3)\cdots(a+r-1)\left(\frac{z}{a}\right)^r$$

we get the series for the exponential:

$$\exp(z) = 1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\cdots$$

Since the ratio of the $(n+1)$ th to the n th term, $\frac{z}{n+1}$, vanishes as $n \to \infty$, this series converges everywhere.

11.7.10 The differential equation

$$\frac{d}{dz}\exp(z) = \exp(z)$$

follows by term by term differentiation.

11.8 The binomial theorem:

$$F(a;z) = (1-z)^{-a} = 1+\frac{a}{1!}z+\frac{a(a+1)}{2!}z^2+\frac{a(a+1)(a+2)}{3!}z^3+\cdots$$

11.9 We note some useful identities in manipulating infinite series.

$$\frac{d}{dz}\sum_{r=0}^{\infty} f_r\frac{z^r}{r!} = \sum_{r=0}^{\infty} f_{r+1}\frac{z^r}{r!}, \qquad z\sum_{r=0}^{\infty} f_r\frac{z^r}{r!} = \sum_{r=0}^{\infty} r f_{r-1}\frac{z^r}{r!}$$

$$z\frac{d}{dz}\sum_{r=0}^{\infty} f_r\frac{z^r}{r!} = \sum_{r=0}^{\infty} r f_r\frac{z^r}{r!}, \qquad z\frac{d^2}{dz^2}\sum_{r=0}^{\infty} f_r\frac{z^r}{r!} = \sum_{r=0}^{\infty} r f_{r+1}\frac{z^r}{r!}$$

11.9.1 Some we have used already others will be useful later. The identities give another proof of the binomial theorem using differential equations.

11.9.2 The the recursion relations

$$(1-z)F(a;z) = F(a-1;z).$$

and

$$F'(a;z) = aF(a+1;z).$$

follow using the above identities.

11.9.3 Combining the two we get the differential equation

$$(1-z)F'(a;z) = aF(a;z)$$

whose solution satisfying $F(a;1) = 1$ is $F(a;z) = (1-z)^{-a}$.

11.9.4 What if we modified the exponential series by putting some factors in the denominator?

11.9.5 **Exercise**:Express the series

$$F_1(c;z) = 1 + \frac{1}{c}z + \frac{1}{c(c+1)}\frac{z^2}{2!} + \frac{1}{c(c+1)(c+2)}\frac{z^3}{3!} + \cdots$$

in terms of Bessel functions. (**Hint**: derive a differential equation for it using the above formulae.)

11.10 The *Confluent Hypergeometric Series* is defined by

$$F_1^1(a,c;z) = 1 + \frac{a}{c}z + \frac{a(a+1)}{c(c+1)}\frac{z^2}{2!} + \frac{a(a+1)(a+2)}{c(c+1)(c+2)}\frac{z^3}{3!} + \cdots$$

11.10.1 For c not equal to a negative integer, this is an entire function of z .

11.10.2 The previous series are special cases:

$$F_1^1(a,a;z) = e^z, \quad \lim_{c\to\infty} F_1^1(a,c;cz) = (1-z)^{-a}.$$

11.10.3 When $a = -n$ is a negative integer, the series terminates and becomes just a polynomial:

$$F_1^1(-n, c; z) = 1 - \frac{n}{c}z + \frac{n(n-1)}{c(c+1)}\frac{z^2}{2!} - \frac{n(n-1)(n-2)}{c(c+1)(c+2)}\frac{z^3}{3!} + \cdots + \frac{(-1)^n}{c(c+1)(c+2)\cdots(c+n-$$

11.10.4 The coefficients satisfy the recursion relation

$$\frac{f_{r+1}}{f_r} = \frac{a+r}{c+r}, \Rightarrow (c+r)f_{r+1} = (a+r)f_r$$

which implies that the function $F_1^1(a, c; z)$ satisfies the following differential equation.

11.11 The *confluent hypergeometric differential equation* is

$$zu'' + (c-z)u' - au = 0.$$

11.11.1 In other words, the Confluent Hypergeometric function is the solution of this differential equation with initial conditions $u(0) = 1, u'(0) = \frac{a}{c}$.

11.11.2 As the order of z grows, the coefficients tend to that of the exponential series; so this entire function tends to e^z for large $|z|$. This can also be seen from the differential equation.

11.11.3 Exercise: Derive an integral representation of the form $F_1^1(a, c; z) = \int_C e^{tz}v(t)dt$ for some function $v(t)$ and a suitable contour on the t -plane.**Hint**: derive a first order differential equation for $v(t)$, solve it and then study which contour will make the integral converge. This is called Laplace's method.

11.12 An equivalent form is the *Whittaker equation*

$$W'' + \left(-\frac{1}{4} + \frac{\kappa}{z} + \frac{\frac{1}{4} - \mu^2}{z^2}\right)W = 0.$$

The change of variable $W = e^{-\frac{z}{2}}z^{\mu+\frac{1}{2}}F(z)$ will take this to the confluent hypergeometric equation.

11.2 First Order Systems

11.13 Any two dimensional system of ordinary differential equations of the form

$$\frac{d\psi}{dz} + A(z)\psi = 0, \quad A(z) = B + \frac{C}{z}$$

can be reduced to the confluent hypergeometric equations.

11.13.1 If $\psi \mapsto z^a e^{bz}\psi$ for constants a, b,

$$A(z) \mapsto A(z) - \frac{a}{z} - b.$$

We can choose $a = \frac{1}{2}\,\mathrm{tr}\,C$ and $b = \frac{1}{2}\,\mathrm{tr}\,B$ such that B, C become traceless after the transformation. In the case of the Coulomb-Dirac equation they are already traceless.

11.13.2 Recall that the square of any traceless 2×2 matrix is a multiple of the identity:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11} \end{pmatrix} = \left(A_{11}^2 + A_{12}A_{21}\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -\det A.$$

11.13.3 Make the substitution

$$\psi = \left(\frac{d}{dz} - A\right)\chi$$

so that the equation becomes

$$\left(\frac{d^2}{dz^2} - A^2 - \frac{dA}{dz}\right)\chi = 0$$

which is

$$\left(\frac{d^2}{dz^2} - A^2(z) + \frac{1}{z^2}C\right)\chi = 0$$

If $C = S^{-1}\begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix} S$ and $\phi = S\chi$ we get the decoupled equations

$$\left(\frac{d^2}{dz^2} - A^2(z) + \frac{1}{z^2}\begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix}\right)\phi = 0.$$

Note also that

$$A^2(z) = \frac{1}{z^2}\left(C_{11}^2 + C_{12}C_{21}\right) + \frac{1}{z}\left(2B_{11}C_{11} + B_{21}C_{12} + B_{12}C_{21}\right) + B_{11}^2 + B_{12}B_{21}$$

Moreover, $c^2 = C_{11} + C_{12}C_{21}$. Thus the equation for ϕ becomes

$$-\phi_1''(z) + \left[\frac{c(c-1)}{z^2} + \frac{1}{z}\left(2B_{11}C_{11} + B_{21}C_{12} + B_{12}C_{21}\right) + \left(B_{11}^2 + B_{12}B_{21}\right)\right]\phi_1 = 0.$$

$$-\phi_2''(z) + \left[\frac{c(c+1)}{z^2} + \frac{1}{z}\left(2B_{11}C_{11} + B_{21}C_{12} + B_{12}C_{21}\right) + \left(B_{11}^2 + B_{12}B_{21}\right)\right]\phi_2 = 0.$$

This can be turned into the Whittaker equation and solved in terms of the confluent hypergeometric form using the transformation mentioned earlier.

11.3 Reduction of the Coulomb-Dirac Equation

11.13.4 The matrices can be read off

$$B(z) = \begin{pmatrix} 0 & m+E \\ m-E & 0 \end{pmatrix} \quad C = \begin{pmatrix} k & Z\alpha \\ -Z\alpha & -k \end{pmatrix}$$

so that

$$c^2 = k^2 - (Z\alpha)^2,$$

$$2B_{11}C_{11} + B_{21}C_{12} + B_{12}C_{21} = -2Z\alpha E,$$

and

$$B_{11}^2 + B_{12}B_{21} = m^2 - E^2$$

so that

$$-\phi_2'' + \left[\frac{c(c+1)}{r^2} - \frac{2Z\alpha E}{r} + m^2 - E^2\right]\phi_2 = 0.$$

This is similar to the Schrödinger equation for the non-relativistic hydrogen atom, but with constants redefined. The main difference is that c may not be an integer.

Chapter 12

Spinorial Harmonics

12.1 The operator $K = 1 + \boldsymbol{\sigma} \cdot \mathbf{L}$ can be thought of as the Dirac operator on the sphere S^2 .

12.1.1 It is the angular part of the Dirac operator on R^3 as we saw in an earlier chapter.

12.2 It can be expressed in terms of the total angular momentum operator

$$K = \mathbf{J}^2 - \mathbf{L}^2 + \frac{1}{4}.$$

where

$$\mathbf{J} = \mathbf{L} + \frac{1}{2}\sigma.$$

12.3 J^2, L^2, J_3 form a set of mutually commuting operators. The eigenvalues of L^2 are $l(l+1)$ for $l = 0, 1, 2, 3, \cdots$ and those of J^2 are $j(j+1)$, with $j = l + \frac{1}{2}$ for $l = 0, 1, 2, \cdots$ and $j = l - \frac{1}{2}$ for $l = 1, 2, 3 \cdots$. For given j , J_3 has eigenvalues $-j, -j+1, \cdots j-1, j$.

12.4 The eigenvalues k of K, J are therefore

$$k = l, \text{for } l = 0, 1, 2, \cdots$$

and

$$K = -l, \text{for } l = 1, 2, \cdots$$

12.4.1 Thus the spectrum of K consists of the integers $k \neq 0$. More precisely, the eigenvalues of K are integers k with degeneracy $2|k|$.

12.5 It is possible to find the eigenspinors by solving the equations $[\sigma \cdot L + 1]Y = kY$ using separation of variables in the angular co-ordinates.

Chapter 13

Relativistic Quantum Statistical Mechanics

13.1 Dirac postulates that the vacuum of the system already has all negative energy states occupied. Also each state can be occupied by at most one particle (exclusion principle).

13.1.1 It is impossible to make sense of relativistic quantum mechanics if we study a single particle, because of the negative eigenvalues of the Dirac operator of a single particle, it cannot describe the total energy.

13.1.2 We are forced to study states containing an arbitrary number of particles and holes: only the difference between the number of particles and total number of holes is conserved. This *particle number* can take any integer value.

13.2 Define the thermodynamic partition function to be

$$Z(q,z) = \sum_{\text{all multi-particle states}} q^{\text{total energy}} z^{\text{particle number}}.$$

The variable $q = e^{-\beta}$ (β is the inverse of temperature) is the Boltzmann factor and z is the fugacity in the language of thermodynamics.

13.3 Suppose the Dirac operator has only discrete eigenvalues λ each with degeneracy $d(\lambda)$. Then the partition function is

$$Z(q,z) = \prod_{\lambda} \left(1 + q^{|\lambda|} z^{\operatorname{sgn}(\lambda)}\right)^{d(\lambda)}.$$

13.3.1 If the Dirac operator has a zero eigenvalue (as for the case of periodic boundary condition) there are several state in the multi-particle system with zero energy: depending on whether the null eigenspace of the Dirac operator is occupied or not.

13.3.2 It is useful to study the Dirac operator in the simplest of all spaces- the circle- as a way to understand the hole theory of Dirac better.

13.3.3 The Dirac algebra is just $\gamma_0^1 = -\gamma_1^1 = 1, \gamma_0\gamma_1 = -\gamma_1\gamma_0$; let $\alpha = \gamma_0\gamma_1$ so that $\alpha^2 = 1$. The eigenvalue problem for energy is

$$-i\alpha\frac{d}{dx}\psi + \beta m\psi = E\psi$$

We can choose a basis with $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .If the mass is zero, we can restrict to a one-component spinor and the eigenvalue equation Weyl operator becomes just

$$-i\frac{d}{dx}\psi = E\psi$$

13.3.4 There are two possible boundary conditions: periodic or anti-periodic

$$\psi(x + 2\pi) = \pm\psi(x).$$

Anti-periodic boundary conditions are allowed because all observables such as energy or electric charge density are quadratic functions of ψ .

13.4 For the periodic (respectively anti-periodic) boundary conditions on the circle, the spectrum of this Weyl operator is the set of integers (respectively half-integers), each with degeneracy one.

13.4.1 Let us consider first the case of anti-periodic boundary conditions, which is simpler because the ground state is unique.

13.4.2 If the spectrum of the Dirac operator does not contain zero, any departure from this state will cost a positive amount of energy. The simplest such state has a particle of energy $\nu > 0$; or a 'hole' corresponding to the eigenvalue $\mu < 0$, which has energy $-\mu$ .We can also have a pair of particles $0 < \nu_1 < \nu_2$ which has energy $\nu_1 + \nu_2$ or a particle-hole pair $\mu < 0 < \nu$ which has the positive energy $\nu - \mu$. The particle number (the number of particles minus the number of anti-particles) is 2 for the first case and 0 for the second.

13.4.3 Thus an arbitrary state of the system is given by a list of unequal positive half-integers and another list of unequal negative half-integers: $\mu_1 < \mu_2 \cdots < \mu_r < 0 < \nu_1 < \nu_2 < \cdots \nu_s$. The corresponding state has 'holes' at $\mu_1 < \mu_2 \cdots < \mu_r$ and particles at $\nu_1 < \nu_2 < \cdots \nu_s$. The total energy is

$$E = \nu_1 + \nu_2 + \cdots + \nu_s - (\mu_1 + \mu_2 + \cdots + \mu_r)$$

which is a positive number. The particle number is $r - s$ which can be positive, zero or negative.

13.5 The partition function for anti-periodic boundary conditions is

$$Z_1(q, z) = \prod_{n=1}^{\infty} \left(1 + q^{n-\frac{1}{2}} z\right) \left(1 + q^{n-\frac{1}{2}} z^{-1}\right)$$

and for periodic boundary condition is

$$Z_0(q, z) = (1 + z) \prod_{n=1}^{\infty} (1 + q^n z) \left(1 + q^n z^{-1}\right)$$

13.5.1 The first factor is for the zero eigenvalue of the Dirac operator.

13.5.2 These functions have remarkable double periodity properties and are related to elliptic functions (the θ functions of Jacobi).

13.6 The Dirac operator on S^2 , $\boldsymbol{\sigma} \cdot \mathbf{L}$, has integer eigenvalues k with degeneracy $2|k + 1|$.

13.6.1 Exercise: Find the partition function of this system.

$$Z(q, z) = \prod_{k=0}^{\infty} \left\{\left(1 + q^k z\right) \left(1 + q^{k+2} z^{-1}\right)\right\}^{2(k+1)}$$

Chapter 14

The Relativistic Degenerate Fermi Gas

14.1 A system is *degenerate* if the mean distance between particles is comparable to the de Broglie wavelength.

14.1.1 This means quantum effects cannot be ignored.

14.2 The energy density P and number density n are related to each other by

$$s\int_{|\mathbf{p}|<p_F}\frac{d^3p}{(2\pi)^3}=n, \quad s\int_{|\mathbf{p}|<p_F}\omega_{\mathbf{p}}\frac{d^3\mathbf{p}}{(2\pi)^3}=P$$

where $\omega_{\mathbf{p}}$ is the energy of a particle of momentum $\mathbf{p}$. Also, s is the number of independent polarization states of a particle with given momentum.

14.2.1 The kinetic energy density of a gas is also its pressure; hence the notation P for it.

14.2.2 For an electron $s=2$, $\omega_{\mathbf{p}}=\sqrt{\mathbf{p}^2+m^2}$.

14.2.3 If the density is large enough, the mean momentum can get large compared to the mass. Then $\omega_{\mathbf{p}}=|\mathbf{p}|$; We can perform the integrals to get

$$s\frac{1}{(2\pi)^3}4\pi\frac{p_F^3}{3}=n, \quad s\frac{1}{(2\pi)^3}4\pi\frac{p_F^4}{4}=P, \Rightarrow P=Cn^{\frac{4}{3}}, \quad C=\frac{1}{8}\left[\frac{6^4\pi^2}{s}\right]^{\frac{1}{3}}.$$

14.3 The density of certain stars, *white dwarfs* is so high that the electrons in them become a relativistic degenerate gas.

14.3.1 The pressure of the electron gas is balanced by the gravitational attraction of the protons.

14.3.2 The protons and electron densities are equal for electrical neutrality; but the protons being much heavier dominate the mass while the electrons dominate the pressure since their de Broglie wavelengths are much longer.

14.4 The energy of the star is given by

$$E = -Gm_p^2 \int \frac{n(x)n(y)}{|x-y|} dxdy + C\int n(x)^{\frac{4}{3}} dx, N = \int n(x)dx.$$

14.4.1 m_p is the mass of the proton and G is Newton's constant in units where $\hbar = c = 1$. In fact in these units $G = m_P{}^{-2}$, the inverse square of the Planck mass so that $\sqrt{G}m_p \sim 10^{19}$.

14.4.2 Here we are assuming that over one de Broglie wavelength of the electron, the gravitational potential is approximately constant: the Thomas-Fermi approximation. We will see that this is easily satisfied as the radius of the degenerate core of the star is several thousand kilometers.

14.4.3 We can factor out the total number of particles by defining $\nu(x) = \frac{n(x)}{N}$:

$$\int \nu(x)dx = 1$$

14.4.4

$$E = -Gm_p^2 N^2 \int \frac{\nu(x)\nu(y)}{|x-y|} dxdy + CN^{\frac{4}{3}} \int \nu(x)^{\frac{4}{3}} dx.$$

$$\frac{E}{N^{\frac{4}{3}}} = -g^2 \int \frac{\nu(x)\nu(y)}{|x-y|} dxdy + C\int \nu(x)^{\frac{4}{3}} dx, \quad g = \sqrt{G}m_p N^{\frac{1}{3}}$$

Here g is a dimensionless quantity: each term has dimensions of $\frac{1}{\text{length}}$.

14.5 There is a constant S such that

$$\int \frac{\nu(x)\nu(y)}{|x-y|} dx dy \geq S \int \nu(x)^{\frac{4}{3}} dx$$

whenever $\int \nu(x) dx = 1, \nu(x) \geq 0$ **.**

14.5.1 The best such constant is the solution of the variational problem

$$S = \inf_{\nu(x)\geq 0, \int \nu(x)=1} \left[\frac{\int \frac{\nu(x)\nu(y)}{|x-y|} dx dy}{\int \nu(x)^{\frac{4}{3}} dx} \right]$$

This becomes an integral equation

$$2 \int \frac{\nu(y)}{|x-y|} dy + \mu\nu(x) = \frac{4}{3} S \nu^{\frac{1}{3}}$$

One can reduce this to an integral equation in one dimension by assuming spherical symmetry.

14.5.2 Exercise:For astrophysical purposes, even a crude estimate based on a sphere of constant number density at the degenerate core is useful. Determine this estimate of S .

14.6 Thus if

$$g > g_c = \sqrt{\frac{C}{S}}$$

the energy diverges and star collapses.

14.7 This gives a bound

$$N_c < \left(\frac{m_P}{m_p} g_c \right)^3$$

for the maximum number of protons that the core can support. The maximum mass of the white dwarf is thus

$$M_c = N_c m_p.$$

14.7.1 This is about 1.4 Solar masses. Indeed, white dwarfs are found to have masses less than this limit. Stars with masses greater than this will explode in a supernova and form an even denser object, a *neutron star*. A neutron star is a degenerate fermi liquid of neutrons and there is a maximum mass for them also. It is harder to calculate this because of the effect of the nuclear force (strong force). If the mass exceeds this limit the star will collapse into a blackhole.

14.7.2 Exercise Determine the critical mass using the estimate for S in the exercise above.

Chapter 15

Particles as Oscillators

15.1 A system of free bosons is equivalent to a harmonic oscillator; to each allowed single-particle state, corresponds one degree of freedom of the harmonic oscillator.

15.1.1 Suppose a single boson can exist in states with energies ω_k for some range of values of k . The energy of a two boson state will be $\omega_k + \omega_{k'}$, of a three boson system $\omega_k + \omega_{k'} + \omega_{k''}$ and so on. The energies add because they are free particles: no interaction energy.

15.1.2 The most general bosonic state is given by the number $n_k = 0, 1, 2, \cdots$ of bosons occupying the k th state. Its energy is $\sum_k n_k \omega_k$.

15.1.3 This happens to be the spectrum of a harmonic oscillator whose characteristic frequencies are ω_k . Thus the hamiltonian of such a system of bosons is

$$H = \frac{1}{2} \sum_k \left[p_k^2 + \omega_k^2 q_k^2 - \omega_k \right]$$

We subtract the constant ω_k to make the ground state energy equal to zero.

15.1.4 With

$$a_k = \frac{\omega_k q_k + i p_k}{\sqrt{[2\omega_k]}} \quad a_k^\dagger = \frac{\omega_k q_k - i p_k}{\sqrt{[2\omega_k]}}$$

we have

$$[a_k, a^\dagger_{k'}] = \delta_{kk'}, \quad [a_k, a_{k'}] = [a^\dagger_k, a^\dagger_{k'}] = 0.$$

and

$$H = \sum_k \omega_k a_k a^\dagger_k.$$

15.2 The Heisenberg equations of motion of a system of bosons are thus

$$\ddot{q}_k + \omega_k^2 q_k = 0.$$

We are to think of q_k as an operator at each time t .

15.2.1 For a system of free bosons of mass m , $\omega_{\mathbf{p}} = \sqrt{[}\mathbf{p}^2 + m^2]$ where $\mathbf{p}$ is the momentum. Thus the equations of motion become

$$\ddot{q}_{\mathbf{p}} + (\mathbf{p}^2 + m^2) q_{\mathbf{p}} = 0.$$

15.3 If we define

$$\phi(t, \mathbf{x}) = \int q_{\mathbf{p}}(t) e^{i\mathbf{p}\cdot\mathbf{x}} \frac{d\mathbf{p}}{(2\pi)^3}$$

we get

$$\ddot{\phi} - \nabla^2\phi + m^2\phi = 0.$$

15.3.1 This is just the Klein-Gordon equation. But there is an important difference: $\phi(x)$ is now an **operator**. Its Fourier components satisfy the commutation relations

$$[q_{\mathbf{p}}, \dot{q}_{\mathbf{p}'}] = i(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}'), \quad [q_{\mathbf{p}}, q_{\mathbf{p}'}] = 0 = [\dot{q}_{\mathbf{p}}, \dot{q}_{\mathbf{p}'}]$$

at equal time. These translate in position space to

$$[\phi(t, \mathbf{x}), \dot{\phi}(t, \mathbf{x}')] = i\delta(\mathbf{x} - \mathbf{x}'), \quad [\phi(t, \mathbf{x}), \phi(t, \mathbf{x}')] = 0 = [\dot{\phi}(t, \mathbf{x}), \dot{\phi}(t, \mathbf{x}')].$$

15.3.2 Thus $\phi(t, \mathbf{x})$ is a **quantum field**, an operator at each point of space-time. When we originally discussed the Kelin-Gordon equation we were thinking of $\phi(t, \mathbf{x})$ as a complex number at each point of space-time. But that only describes a single particle at a time. Now our system can have as many particles as the available energy allows.

15.3.3 The energy is the operator

$$H = \frac{1}{2}\int [\dot{\phi}^2 + |\nabla\phi|^2 + m^2\phi^2] d\mathbf{x}$$

up to an overall additive constant.

Chapter 16

Fermionic Oscillators

16.1 The state of a system of identical fermions is anti-symmetric under the interchange of a pair of particles.

16.1.1 In particular if two particles have the same quantum numbers, the state would vanish: the Pauli exclusion principle that each single particle state can be occupied .

16.1.2 Consider a system of identical free fermions with single particle states of energy ω_k labelled by k . Each state can be either occupied or not. So the energy is $\sum_k \nu_k \omega_k$, for occupation numbers $\nu_k = 0, 1$.

16.1.3 Let $|0>$ be the empty state, $b_k^\dagger$ the operator that creates a fermion at state k and b_k the one that annihilates such a fermion. Then it is reasonable that $b_k|0>= 0, b_k^{\dagger\ 2} = 0 = b_k^2.$

16.1.4 Also $b_k^\dagger b_{k'}^\dagger|0>= -b_{k'}^\dagger b_k^\dagger|0>$ by anti-symmetry of states.

16.2 The Canonical Anti-Commutation Relations

$$[b_k, b_{k'}^\dagger]_+ = \delta_{kk'}, \quad [b_k, b_{k'}]_+ = 0 = [b_k^\dagger, b_{k'}^\dagger]_+ = 0$$

are satisfied by the creation-annihilation operators of a system of fermions. The empty state ('vacuum') is defined by

$$b_k|0>= 0.$$

The remaining states are built from it by the action of the creation operators

$$b_k^\dagger|0>, b_k^\dagger b_{k'}^\dagger|0>, \cdots$$

In general,

$$b_{k_1}^\dagger b_{k_2}^\dagger \cdots b_{k_r}^\dagger|0>, k_1 < k_2 < k_2 \cdots k_r.$$

form an orthonormal basis of states.

16.2.1 The number operator

$$b_k^\dagger b_k$$

has square zero; hence its eigenvalues are $0, 1$. The hamiltonian of a system of free fermions is

$$H = \sum_k \omega_k b_k^\dagger b_k.$$

16.2.2 The Heisenberg equations of motion

$$\frac{db_k}{dt} = i[H, b_k], \quad \frac{db_k^\dagger}{dt} = i[H, b_k^\dagger]$$

become

$$\frac{db_k}{dt} = -i\omega_k b_k, \quad \frac{db_k^\dagger}{dt} = i\omega_k b_k^\dagger.$$

16.2.3 Unlike in the bosonic case, it is not possible to represent the fermionic creation-annihilation operators as differential operators on functions of real or complex variables. But we can invent a new number system, the *Grassmann numbers* that make this possible. Since about half of all elementary particles are fermions, this number system must be as important as real or complex numbers, although humans only invented them fifty years ago.

16.2.4 Let us begin with a fermionic system with just one single particle state. So there are two possible states altogether: either it is occupied or it is empty.

16.3 A Grassmann number θ satisfies

$$\theta^2 = 0.$$

A complex-valued function of a Grassmann variable is

$$f(\theta) = f_0 + f_1\theta$$

The product and sum of such functions are defined as below:

$$[f+g](\theta) = (f_0+g_0) + (f_1+g_1)\theta, \quad fg(\theta) = f_0g_0 + (f_0g_1 + f_1g_0)\theta.$$

16.3.1 Thus

$$\theta f(\theta) = f_0\theta, \quad \frac{\partial}{\partial\theta} f(\theta) = f_1.$$

16.3.2 Also,

$$\theta^2 f = 0 = \frac{\partial^2 f}{\theta^2}, \quad \frac{\partial(\theta f)}{\partial\theta} + \theta\frac{\partial f}{\partial\theta} = f.$$

In other words, these operators satisfy the relations for fermion creation-annihilation operators:

$$\theta^2 = 0 = \frac{\partial^2}{\partial\theta^2} \quad [\theta, \frac{\partial}{\theta}]_+ = 1.$$

16.3.3 When there are several such variables, we postulate

$$\theta^i\theta^j + \theta^j\theta^i = 0.$$

The most general function is then

$$f(\theta) = f_0 + f_i\theta^i + f_{ij}\frac{\theta^i\theta^j}{2!} + \cdots f_{i_1\cdots i_r}\frac{\theta^{i_1}\cdots\theta^{i_r}}{r!} + \cdots$$

where the coefficients are anti-symmetric

$$f_{ij} = -f_{ji}, \quad f_{i_1\cdots i_r} = \operatorname{sgn}(\pi) f_{i_{\pi(1)}\cdots i_{\pi(r)}}$$

for any permutation π . The coefficient $f_{i_1\cdots i_r}$ can be thought of as the wavefunction of a system of r identical fermions.

16.4 The differentiation of a function of Grassmannian variables is defined as

$$\frac{\partial}{\partial\theta_i} f = f_i + f_{ij}\theta^j + \cdots f_{ii_1\cdots i_r}\frac{\theta^{i_1}\cdots\theta^{i_r}}{r!} + \cdots$$

The multiplication and differentiation operators satisfy

$$[\theta^i, \theta^j]_+ = 0 = [\frac{\partial}{\partial\theta_i}, \frac{\partial}{\partial\theta_j}]_+, \quad [\frac{\partial}{\partial\theta_i}, \theta_j]_+ = \delta_i^j.$$

16.4.1 These are just the creation-annihilation operators of fermions:

$$b_k = \frac{\partial}{\partial\theta^k}, \quad b^{k\dagger} = \theta^k.$$

Thus the hamiltonian of free fermions is the operator

$$H = \sum_k \omega_k \theta^k \frac{\partial}{\partial\theta_k}.$$

Chapter 17

The Two Level Atom

17.1 A simple model of an electron interacting with a photon is

$$H = \omega b^\dagger b + \epsilon_1 a_1^\dagger a_1 + \epsilon_2 a_2^\dagger a_2 + g b^\dagger a_1^\dagger a_2 + g^* b a_2^\dagger a_1$$

17.1.1 Here $b, b^\dagger$ are the creation operators for the boson, $a, a^\dagger$ those for the fermion:

$$[b, b^\dagger] = 1, \quad [a_a, a_b]_+ = 0 = [a_a^\dagger, a_b^\dagger]_+, \quad [a_a, a^\dagger]_+ = \delta_{ab}.$$

17.1.2 This describes the emission of a boson while the fermion makes a transition from the excited state to the ground state or the absorption during the opposite process.

17.2 The states containing a fixed number of bosons, and either both fermionic states occupied or both empty are eigenstates of this hamiltonian; so also when the electron in the lower level with no photon present.

17.2.1 In this case the interaction terms vanish identically.

17.2.2 The most general state state containing one electron is

$$\sum_{n=0}^{\infty} \chi_n |n, 1, 0> + \sum_{n=0}^{\infty} \phi_n |n, 0, 1> .$$

The eigenvalue equation for the hamiltonian becomes

$$\sum_{n=0}^{\infty}(n\omega+\epsilon_1)\chi_n|n,1,0>+\sum_{n=0}^{\infty}(n\omega+\epsilon_2)\phi_n|n,0,1>+$$
$$g\sum_{n=0}^{\infty}\sqrt{n+1}\phi_n|n+1,1,0>+g^*\sum_{n=0}^{\infty}\sqrt{n}\chi_n|n-1,0,1>=$$
$$E\left[\sum_{n=0}^{\infty}\chi_n|n,1,0>+\sum_{n=0}^{\infty}\phi_n|n,0,1>\right]$$

17.2.3 This becomes the recursion relations

$$(n\omega+\epsilon_1)\chi_n+g\sqrt{n}\phi_{n-1}=E\chi_n$$

$$(n\omega+\epsilon_2)\phi_n+g^*\sqrt{n+1}\chi_{n+1}=E\phi_n$$

Shifting the first equation by one step, we get the 2×2 system

$$\begin{pmatrix}(n+1)\omega+\epsilon_1 & g\sqrt{n+1}\\ g^*\sqrt{n+1} & n\omega+\epsilon_2\end{pmatrix}\begin{pmatrix}\chi_{n+1}\\ \phi_n\end{pmatrix}=E\begin{pmatrix}\chi_{n+1}\\ \phi_n\end{pmatrix}$$

Setting the determinant to zero gives the eigenvalue equation

$$E^2-[e_1+e_2]\,E+e_1e_2-|g|^2(n+1)=0,\quad e_1=(n+1)\omega+\epsilon_1,\quad e_2=n\omega+\epsilon_2.$$

17.3 The energy eigenvalues are, for $n=0,1,2,\cdots$

$$E_{n\pm}=(n+\frac{1}{2})\omega+\frac{1}{2}[\epsilon_2+\epsilon_1]\pm\frac{1}{2}\sqrt{(\epsilon_2-\epsilon_1-\omega)^2+4(n+1)|g|^2}.$$

17.3.1 Thus for each value of n there are two eigenstates; the splitting between them ,

$$\sqrt{(\epsilon_2-\epsilon_1-\omega)^2+4(n+1)|g|^2},$$

is increased by the interaction. The absorption and re-emission of photons has caused this change in the energy levels. For each n , the eigenstate is a linear combination of an n photon state with the electron in its upper level and an $n+1$ photon state with the electron in the lower level. The case of 'resonance' is $\epsilon_2 = \epsilon_1 + \omega$ when these two states are degenerate in the absence of interaction; the interaction has the biggest effect in this case by removing the degeneracy.

17.3.2 Exercise: Determine the corresponding eigenvector.

17.3.3 The ground state of an electron is just the state $|0, 1, 0>$, where it occupies the lower level with no photon present.This is an exact eigenstate. The first excited state is when it is in the upper level with no photon present; but this is not an exact eigenstate. A linear combination of this state with the one where the electron is in the lower level with a photon present is the true eigenstate- the case $n = 0$ above. This can be understood as the result of repeated emission and absorption of the photon by the electron.

17.4 This model is often used to describe the interaction of an atom with a laser.

Chapter 18

The Lee Model

18.1 T. D. Lee proposed a model for renormalization:

$$H = \int \frac{dk}{(2\pi)^3} \phi^\dagger(k)\phi(k)\omega_k + \epsilon_1 a_1^\dagger a_1 + \epsilon_2 a_2^\dagger a_2 + g\phi_0^\dagger a_1^\dagger a_2 + g^* \phi_0 a_2^\dagger a_1$$

18.1.1 Here $\phi, \phi^\dagger$ are bosonic creation-annihilation operators and $a, a^\dagger$ fermionic ones:

$$[\phi(k), \phi^\dagger(k')] = (2\pi)^3 \delta(k - k'), \quad [a_\alpha, a_\beta^\dagger]_+ = \delta_{\alpha\beta}$$

The remaining pairs of opertaors commute.

18.1.2 Also

$$\phi_0 = \int \phi(k) \frac{dk}{(2\pi)^3}, \quad \phi_0^\dagger = \int \phi^\dagger(k) \frac{dk}{(2\pi)^3}$$

This is the value of the bosonic field at the origin in position space.

18.1.3 The energy of the boson can be taken to be $\omega_k = \sqrt{[k^2 + m^2]}$ or $\omega_k = \frac{k^2}{2m} + m$ in the non-relativistic approximation.

18.1.4 Thus we have a boson interacting with a fermion. The fermion is allowed to sit only at one point, the origin of position space. It can exist in one of two states, with energies (masses) $\epsilon_{1,2}$. Thus this is much like the 'two level atom' except that the boson is allowed to have many possible energies and momentum. The idea is that the fermion is much heavier than the boson; that is why we can ignore the motion of the fermion after it emits or absorbs a boson. Momentum is not conserved in this approximation, but energy is.

18.1.5 The original purpose of the model was to explain the interaction of a nucleon with pions. The two states of the nucleon are the neutron and the proton. They are both much heavier ($\sim 10^3$ MeV) than the pion (10^2 Mev). Lee had hoped that the neutron-proton mass difference can be explained this way but because of the divergences of this model, the idea didn't work. But it teaches us how to deal with such divergent quantities, by renormalization. This idea is crucal in modetn particle physics, but in the true context of non-abelian Yang-Mills theories it is much harder to understand. So we practice here.

18.1.6 The total number of fermions

$$a_1^\dagger a_1 + a_2^\dagger a_2$$

is conserved.

18.1.7 The number of bosons

$$N_\phi = \int \phi^\dagger(k)\phi(k)\frac{dk}{(2\pi)^3}$$

not conserved. But the quantity

$$N = N_\phi - a_1^\dagger a_1$$

is conserved. When a boson is emitted, a fermion of type 1 is converted to one of type 2.

18.1.8 The states that contain no fermion are uninteresting; it is just a theory of free bosons.

18.1.9 We seek to calculate the energy difference between the two states of the fermion. This is affected by the emission and reabsorption of the boson.So we look at states that contain one fermion.

18.1.10 The only state with $N = -1$ is the one with a single fermion in the 1 state and no bosons. So this must be an eigenstate of the hamiltonian:

$$H|0>|1>=\epsilon_1|0>|1>.$$

18.1.11 The most general states with $N = 0$ are

$$|u,\psi>=u|0>|2>+|\psi>|1>$$

where

$$|\psi>=\int\psi(k)\phi^\dagger(k)\frac{dk}{(2\pi)^3}|0>$$

Let us look for an eigenstate in this sector.

18.1.12 The equation $H|u,\psi>=E|u,\psi>$ gives

$$[\epsilon_2-E]u+g^*\int\psi(k)\frac{dk}{(2\pi)^3}=0,\quad [\omega_k+\epsilon_1-E]\psi(k)+gu=0$$

wherefrom

$$\psi(k)=\frac{u}{E-[\omega_k+\epsilon_1]}$$

$$E-\epsilon_2=|g|^2\int\frac{1}{E-[\omega_k+\epsilon_1]}\frac{dk}{(2\pi)^3}.$$

18.1.13 Unfortunately this integral is divergent. So we dont get a meaningful physical answer for the energy of the first excited state. This is typical of the problems we encounter in quantum field theory.

18.1.14 The culprit is the integration over all momenta. It is clear that our approximation of no recoil will break down for large enough momenta. So we should only integrate over momenta smaller than spme cut-off Λ which itself is smaller than the mass of the fermions $\Lambda << \epsilon_1, \epsilon_2$. So we go back and replace $\phi_0, \phi_0^\dagger$ in the hamiltonian with

$$\phi_\Lambda = \int_{|k|<\Lambda} \phi(k) \frac{dk}{(2\pi)^3}$$

and its adjoint. Also we make the parameters in the hamiltonian depend on Λ such that physical quantities are independent of Λ.

18.1.15 The energy of the first excited state is a physically meaningful quantity. But the parameters in the hamiltonian don't have to be. ϵ_1 is a physical quantity because it is the energy of the state with $N = -1$. So we don't mess with that. We can make ϵ_2 depend on Λ. This process is called regularization.

$$H_\Lambda = \int \frac{dk}{(2\pi)^3} \phi^\dagger(k)\phi(k)\omega_k + \epsilon_1 a_1^\dagger a_1 + \epsilon_2(\Lambda) a_2^\dagger a_2 + g\phi_\Lambda^\dagger a_1^\dagger a_2 + g^* \phi_\Lambda a_2^\dagger a_1$$

18.1.16 Then we determine its dependence on Λ in order that the physical quantity E_2, the energy of the excited state a finite quantity. This is called renormalization.

18.1.17 In fact

$$\epsilon_2(\Lambda) = E_2 - |g|^2 \int_{|k|<\Lambda} \frac{1}{E_2 - [\omega_k + \epsilon_2]} \frac{dk}{(2\pi)^3}.$$

18.1.18 In effect we have traded the ϵ_2 parameter in the original hamiltonian (the unperturbed first excited state energy) for the exact energy of the first excited state E_2. All physical quantities will be independent of Λ if thought of as functions of E_2 rather than as functions of ϵ_2.

18.1.19 Next we can look for a scattering state for the 'pion' against a static 'proton'. This invloves solving the inhomogenous Schr "odinger equation $[H - E]|u, \psi >= |v, \chi >$ where χ represents an incoming pion state. We will get

$$[\epsilon_2(\Lambda) - E]u + g^* \int_{|k|<\Lambda} \psi(k) \frac{dk}{(2\pi)^3} = v, \quad [\omega_k + \epsilon_1 - E]\psi(k) + gu = \chi(k)$$

18.1.20 We can again eliminate $\psi(k)$:

$$\psi(k) = \frac{gu}{E - [\omega_k + \epsilon_1]} + \frac{\chi(k)}{\omega_k + \epsilon_1 - E}.$$

Also,

$$[\epsilon_2(\Lambda) - E]u - |g|^2 u \int_{|k|<\lambda} \frac{1}{E - [\omega_k + \epsilon_1]} \frac{dk}{(2\pi)^3} = v - g^* \int_{|k|<\Lambda} \frac{\chi(k)}{\omega_k + \epsilon_1 - E} \frac{dk}{(2\pi)^3} + v.$$

18.1.21 Trading $\epsilon_2(\Lambda)$ for the more physical E_2 we get

$$u|g|^2 \int_{|k|<\Lambda} \left[\frac{1}{E_2 - [\omega_k + \epsilon_1]} - \frac{1}{E - [\omega_k + \epsilon_1]} \frac{dk}{(2\pi)^3} \right] = v - g^* \int_{|k|<\Lambda} \frac{\chi(k)}{\omega_k + \epsilon_1 - E} \frac{dk}{(2\pi)^3}.$$

18.1.22 If $m << E_2 - \epsilon_1$ (as is true for the pion and nucleon) we can use the non-relativistic formula for the pion energy $\omega_k = m + \frac{k^2}{2m}$. In this case the integral in the l.h.s. converges. It is possible to determine the scattering phase shift of the pion from the proton from the above equation from the integral equation

$$u|g|^2 \int \left[\frac{1}{E_2 - [\omega_k + \epsilon_1]} - \frac{1}{E - [\omega_k + \epsilon_1]} \frac{dk}{(2\pi)^3} \right] = v - g^* \int \frac{\chi(k)}{\omega_k + \epsilon_1 - E} \frac{dk}{(2\pi)^3}.$$

18.1.23 But if $m << E_2 - \epsilon_1$ we must use the relativistic formula and the integral is still divergent. In this case we must do yet another renormalization, where we allow the 'coupling constant' g to become a function of Λ also. Then we must determine its dependence by fixing some other physical quantity. It turns out that it is possible to do this, but that takes us out of the scope of this course.

18.1.24 The renormalization of the Lee model and its more accurate predecessor, the Yukawa model, is becoming relevant again in the context of the Higgs boson. In this case there is one massive fermion, the top quark. The boson has a mass comparable to, or larger, than the top quark. The remaining fermions are all (in this scale) massless. There are still unsolved problems in the renormalization of this theory, which will be interesting as the LHC starts producing data on the Higgs. Stay tuned.